



Establishing a Framework for Evaluating Teaching Excellence: An Overview of the Literature

**Office of Institutional Analysis & Planning
March 2007**

Establishing a Framework for Evaluating Teaching Excellence: An Overview of the Literature

Executive Summary

Within the last twenty years, there has been an increased focus on the importance of an undergraduate education, a reconsideration of the intrinsic value of teaching, and an interest in establishing criteria to identify excellent teaching for both formative and summative evaluation purposes. Many of the early studies on teaching effectiveness were grounded in the behaviorist research tradition that reduced teaching to a set of technical skills; however, more contemporary studies have recognized that teaching is a complex, interactive activity and is therefore influenced by a number of factors. As many researchers have come to recognize, “teaching is more than transmitting knowledge; good teaching transforms people...” (Braskamp, 2000). Other researchers have noted that an ongoing commitment to scholarship and life-long learning as well as participation in the academic community also play a role in teaching excellence.

There is general consensus in the research that teaching competencies can be grouped into three main categories: content expertise, instructional design, and instructional delivery. Much of the research on identifying specific characteristics associated with excellent teaching has been based on analyses of students’ evaluations of their instructors. Although the use of student ratings as an evaluation tool has been quite controversial in the field, meta-analyses of these studies have demonstrated that student feedback provides a valid index—in conjunction with other sources of evaluation data—of instructional effectiveness.

Overall, the research indicates that excellent teaching consists of a number of competencies related to content expertise, instructional design, and instructional delivery, as well as interpersonal traits such as helpfulness, friendliness, and open-mindedness. Effective instructors demonstrate the following characteristics:

- They promote learning by clearly communicating expectations and goals;
- they know their subject matter and are able to relate course material back to real life situations;
- they are enthusiastic and stimulate students’ interest in the subject
- they are organized and prepared for class;
- they develop rapport with their students and show both concern and genuine respect for their students;
- they provide timely feedback and monitor students’ comprehension of the material presented.

In addition to the aforementioned characteristics, the results of a study conducted by Shaw and Young (1999) indicated that excellent instructors consistently received very high ratings on the following three items:

- value of the course
- concern for student learning, and
- genuine respect for students.

In addition to identifying dimensions of excellent teaching, it is also important to develop a comprehensive, valid, and reliable framework for evaluating teaching effectiveness that is perceived as rigorous, credible, and respectful. There is some consensus in the field that evaluation frameworks—comprised of the policies, procedures and criteria used—should be based on the following principles:

- Evaluation should remain neutral with respect to individual style and beliefs, yet acknowledge and respect diversity across faculty and disciplines.
- Evaluation should rest upon multiple and credible sources of data that are valid, consistent and reliable.
- Evaluation should assess both the substantive and technical skills components of teaching.
- Evaluation should contribute to improved teaching practice and career development in general.
- Evaluation should be done in consultation with the key individuals responsible for the process.
- Evaluations should be based on clear communication of expectations by the evaluator(s) and the instructor being evaluated.

The evaluation process is comprised of four major components:

- the range of characteristics evaluated (e.g., content expertise, instructional design, instructional delivery, evidence of commitment to improving teaching practice, participation in the academic community)
- the range of evaluation sources available to supply data (students and alumni, peers, colleagues, Dean, Chair, self)
- the range of data types available (e.g., teaching materials, student ratings, classroom observation, structured interviews, teaching portfolios, reflective statements), and;
- criteria by which the data are to be evaluated (reliability, validity, objectivity of qualitative and quantitative data sources, weighting assigned to various data sources and types).

There is some consensus in the literature that the validity, reliability and usefulness of the data, as well as the fairness and accountability of the evaluation process can be improved by:

- Training evaluation sources to recognize and avoid certain types of rater errors before they undertake their assessment tasks.
- Providing evaluators—whether students or peers—with a structured set of questions with regard to the teaching competencies to be evaluated.
- Standardizing rating procedures and criteria used in an evaluation process.
- Establishing a replicable protocol, particularly with respect to (1) the administration and collection of students' evaluations of instructors and (2) classroom observations by peers or colleagues.

Teaching portfolios and reflective statements are also increasingly incorporated into the evaluation process as valid sources of data supplied by the instructor. As with other data submitted for evaluation purposes, there must be some guidelines to ensure the content, size, and usefulness of such information if it is to be used effectively in evaluating whether an instructor is effective and/or demonstrates excellent teaching practices.

In recent years, there has been a trend toward making evaluation policies and procedures more transparent. Whereas evaluation frameworks were previously embedded in personnel policies and faculty agreements, a number of institutions in both Canada (e.g., University of Toronto, University of Saskatchewan, and University of Alberta) and the United States (e.g. Cornell University, Arizona State University) have moved toward publicizing comprehensive evaluation frameworks on their institutional web sites.

There also seems to be some agreement in the literature that teaching evaluation processes that are not grounded in a philosophy of promoting teaching excellence and/or do not provide professional development opportunities to support improving teaching practice tend to generate both anxiety and skepticism on the part of faculty. A number of Canadian and American universities have developed an extensive and comprehensive set of instructional resources to support faculty in becoming excellent teachers. The teaching centres or services in these institutions also make many of their services and publications available on their web sites. Noteworthy Canadian examples include the University of Manitoba, the University of Saskatchewan and the University of Alberta.

This page intentionally left blank.

Introduction

...traditionally, teaching has generally been accepted by university faculty and administrators as a less valuable activity compared to other aspects of academic life...

...even in colleges and university-colleges that are ostensibly "teaching-led" institutions (1997:24).

In conjunction with both an escalating concern over demonstrating public accountability and an increased focus on the importance of an undergraduate education, there has been a growing trend in higher education toward re-assessing the value of faculty's teaching activities and establishing criteria to determine excellence in teaching...

Many of the early studies on teaching effectiveness were characterized by a narrow focus on observable behaviours and a tendency to reduce teaching to a set of technical skills...

...research into the subject of

A number of researchers in higher education have noted that traditionally, teaching has generally been accepted by university faculty and administrators as a less valuable activity compared to other aspects of academic life such as research (Boyer, 1990, Seldin, 1993, cited in Pratt, 1997; Webbstock, 1999). According to Pratt, a researcher at the University of British Columbia, there has been "a tendency among faculty to subordinate teaching to research" even in colleges and university-colleges that are ostensibly "teaching-led" institutions (1997:24). Pratt maintains that as colleges expanded into university-colleges with degree granting status, they developed hiring policies that have attracted instructors with more advanced degrees (Ph.D.); however, "what may not be considered in the decision to hire [candidates] with more advanced degrees is the enculturation they have gone through ... and the effect it has had on [their] thinking about what kind of activity ... [research over teaching] is most valued" (1997:24).

In conjunction with both an escalating concern over demonstrating public accountability and an increased focus on the importance of an undergraduate education, there has been a growing trend in higher education toward re-assessing the value of faculty's teaching activities and establishing criteria to determine excellence in teaching for both formative and summative evaluations (Cornell Centre for Learning and Teaching, 1997). Yet, despite a growing number of persuasive voices who have argued in favour of recognizing teaching as a scholarly activity and a legitimate route to promotion (Pratt, 1997; Webbstock, 1999), many faculty and administrators have been slow to buy into the belief that teaching is a valued activity. The reluctance may have arisen because faculty did not see this reflected in the reward structures of promotion and tenure in their institutions (Pratt, 1997), or because they were resistant to being subjected to evaluation procedures that failed to recognize the complexity of teaching endeavours (Young & Shaw, 1999; Tigelaar, Dolmans, Wolfhagen & Van der Vleuten, 2004).

Many of the early studies on teaching effectiveness were grounded in the behaviourist research tradition; a linear model characterized by a narrow focus on observable behaviours and a tendency to reduce teaching to a set of technical skills—at the expense of considering the skills required to effectively communicate subject matter to students (Cornell Centre for Learning and Teaching, 1997). It has also been claimed that "teaching is more than transmitting knowledge; good teaching touches and transforms people through relationships and community" (Braskamp, 2000:23). Indeed, research into the subject

teaching effectiveness has increasingly recognized that teaching is a complex activity involving...

Teaching effectiveness is also influenced by a number of other factors...

**Given that instructors in higher education "are, for the most part, educated to be justifiable authorities on the subjects they teach, but [are] only...indirectly educated in how to teach those subjects".
...faculty tend to react with scepticism, mistrust, or anxiety when summative evaluations of their teaching competencies appear to be based primarily on ratings of their classroom actions or technical skills, whether by their students or their peers.**

of teaching effectiveness has increasingly recognized that teaching is a complex activity involving instructors' thought processes—including beliefs, intentions, and neurolinguistic programming that influences preferred modes of teaching and learning (Pratt, 1997; Brown, 2004). Teaching effectiveness is also influenced by a number of other factors: the context in which an individual teaches (e.g., institutional mission statements, resources, support for professional development, type of course) and changing visions of student learning and the role of the instructor (Young & Shaw, 1999; Tigelaar, Dolmans, Wolfhagen & Van der Vleuten, 2004).

Given that instructors in higher education "are, for the most part, educated to be justifiable authorities on the subjects they teach, but [are] only...indirectly educated in how to teach those subjects" (Cornell Centre for Learning and Teaching, 1997:12), faculty tend to react with scepticism, mistrust, or anxiety when summative evaluations of their teaching competencies appear to be based primarily on ratings of their classroom actions or technical skills, whether by their students or their peers. In order to answer the questions (1) what constitutes excellent teaching, and (2) how might an institution develop an evaluation framework that reflects both a student-centred teaching orientation and the fact that an instructor's teaching competencies involve more than just a set of skills employed in the classroom to impart knowledge and skills to students, the remainder of this literature review will (1) identify the activities and thought processes that contribute to excellent teaching; (2) consider the main principles that contribute to comprehensive evaluation frameworks and (3) outline several evaluation procedures and modes of data collection that have been incorporated into teaching evaluation processes at other institutions.

Identifying Dimensions of Excellent Teaching

Cornell Centre for Learning and Teaching recommends an integrative model that links teaching back to research and reconciles the dual tasks of both understanding and explaining a body of knowledge. According to this model, "effective teachers are able to understand enough about their students' ways of thinking that they [teachers] can translate their own understanding of subject matter into a form that connects with their students."

...the competencies that have been consistently identified as universal and integral to excellent teaching can be classified into three broad groups: content expertise, instructional design skills... and instructional delivery skills

Some researchers... maintain that evidence of life-long learning and scholarship as well as active participation in academic communities both... should be considered when assessing effective teaching.

...on-going reflection and self-evaluation are also measures of an effective instructor...

The use of student ratings of instructors is controversial within the field...

However, meta-analyses of the studies on student evaluations have demonstrated that students' feedback provides a valid index of instructional effectiveness.

The Centre for Learning and Teaching at Cornell University contends that a narrow focus on either technical process skills or translating abstract information into understandable terms limits the ability to understand what teaching is (1997). In place of a dichotomous approach, the Cornell Centre for Learning and Teaching recommends an integrative model that links teaching back to research and reconciles the dual tasks of both understanding and explaining a body of knowledge. According to this model, "effective teachers are able to understand enough about their students' ways of thinking that they [teachers] can translate their own understanding of subject matter into a form that connects with their students." (1997)

While many researchers have criticized an overly narrow focus on a technical approach to instruction, they do acknowledge that such competencies are important and do contribute to effective teaching (Pratt, 1997). As such, the competencies that have been consistently identified as universal and integral to excellent teaching can be classified into three broad groups: content expertise, instructional design skills (e.g., course design and planning, development of assignments and exams used to evaluate student learning), and instructional delivery skills such as presentation skills and interactions with students both inside and outside of the classroom (Cornell Centre for Learning and Teaching, 1997; Pratt, 1997; Young & Shaw, 1999; Tigelaar et al, 2004). Some researchers (e.g., Pratt, 1997, Braskamp, 2000, Tigelaar et al, 2004) also maintain that evidence of life-long learning and scholarship as well as active participation in academic communities both within and external to the specific institution should be considered when assessing effective teaching. Porter and Brophy contended that on-going reflection and self-evaluation are also measures of an effective instructor (1988; cited in Cornell Centre for Learning and Teaching, 1997).

Much of the research aimed at identifying specific traits or actions associated with effective teaching have relied on analyses of students' evaluations of their instructors and courses. The use of student ratings of instructors is controversial within the field; some authors argue that the validity of students' ratings are influenced by superficial criteria as well as by the students' actual or expected final grade in a course. However, meta-analyses of the studies on student evaluations have demonstrated that students' feedback provides a valid index of instructional effectiveness. As one researcher has pointed out, "students are in class almost every day and they know what's going on...they have some sense of whether they are learning" (McKeachie, 1983:38; cited in

Young & Shaw, 1999).

...general consensus... that effective instructors demonstrate the following characteristics...

There is general consensus in the research (e.g., Feldman, 1976, Marsh, 1987, cited in Young & Shaw, 1999) that effective instructors demonstrate the following characteristics:

- they promote learning by clearly communicating expectations and goals;
- they know their subject matter; they are enthusiastic and stimulate students' interest in the subject;
- they are organized and prepared for class;
- they develop rapport with their students and show both concern and respect for their students; they provide timely feedback and monitor students' comprehension of the material presented, and;
- they relate course material back to real life situations.

While some researchers (Marsh, 1987, cited in Young & Shaw, 1999) have reported that factors such as workload difficulty and examinations/grading are consistently reflected in students' views of effective teaching, other researchers (e.g., Feldman, 1976, 1988; cited in Young & Shaw, 1999) have suggested that characteristics related to classroom management—such as course difficulty and workload—were perceived as less important characteristics of effective teachers than factors such as enthusiasm, certain interpersonal traits, and the instructor's knowledge of the subject. According to Feldman's study, "stimulation of interest and clarity of presentation were the two most highly related dimensions of good teaching" (1976; cited in Young & Shaw, 1999:674). In addition, interpersonal traits such as helpfulness, friendliness and open-mindedness are also considered by students to be important traits of effective teachers (Feldman, 1976; cited in Young & Shaw, 1999:674).

...characteristics related to classroom management—such as course difficulty and workload—were perceived as less important characteristics of effective teachers than factors such as enthusiasm, certain interpersonal traits, and the instructor's knowledge of the subject.

...Young and Shaw's study (1999)... excellent teachers consistently received high ratings on seven items...

The results of Young and Shaw's study (1999) of students asked to rate the effectiveness of a former teacher indicate that excellent teachers consistently received high ratings on seven items: (1) value of the course; (2) motivating students, (3) course organization, (4) effective communications, (5) comfortable learning environment, (6) concern for student learning, and (7) genuine respect for student learning. Furthermore, the instructors rated most effective by students were consistently rated very high on three specific items: value of the course, concern for student learning and genuine respect for students, even if they were ranked slightly lower on the other five items. Several of these items—motivating students by engaging their minds, showing genuine respect for students, and ensuring a

...the instructors rated most effective by students were consistently rated very high on three specific items: value of the course, concern for student learning and genuine respect for students.

good climate for learning—are reflected in the principles of good teaching that inform the University of Alberta’s GFC Policy Manual.

Young and Shaw’s research showed that the value of the course...emerged... strongly as a predictor of teaching effectiveness...

Young and Shaw expressed some surprise that the value of the course had emerged so strongly as a predictor of teaching effectiveness, and hypothesized that as time passed and students had some distance from a course, they may have become more aware of the usefulness or value of the course (1999: 683). Young and Shaw’s observed results, vis-à-vis, the role of “course value” as a predictor of effective teaching would appear to confirm an earlier research study by Nehari and Bender in which the authors hypothesized that the perceived meaningfulness and value of a course were related to students’ perceptions of what they learned in the cognitive-subject matter, affective-personal, and behavioural domains (1978).

... role of “course value” as a predictor of effective teaching...confirms an earlier research study by Nehari and Bender in which the authors hypothesized that the perceived meaningfulness and value of a course were related to students’ perceptions of what they learned...

...students’ and instructors’... conceptions as to the characteristics associated with effective instructors...

In response to challenges that students and instructors may have quite different conceptions as to the characteristics associated with effective instructors, at least two studies have been conducted that examined this issue. Feldman (1988) found that students and instructors were similar, but not identical in their views. Whereas students were likely to place more emphasis on instructional characteristics such as stimulating interest in the subjects, clarity of presentations, availability and helpfulness, faculty were more likely to define effective teaching on the basis of factors such as motivating and setting high standards for students, being intellectually challenging and encouraging self-initiated learning on the part of students (1988:291).

...similar, but not identical...

... students... place more emphasis on instructional characteristics such as stimulating interest... clarity of presentations, availability and helpfulness... faculty were more likely to define effective teaching on the basis of factors such as motivating and setting high standards

...Goldstein and Benassi investigated whether students and instructors held similar beliefs about excellent teaching based on different instructional methods...

A more recent study by Goldstein and Benassi (2006) investigated whether students and instructors held similar beliefs about excellent teaching based on different instructional methods—lecturing compared to discussion leading—and the types of skills associated with them. Based on the results of their studies, Goldstein and Benassi concluded that “to the extent that students and teachers do not agree on the characteristics of ideal lecturers and discussion leaders, their respective views of the classroom experience may differ” (2006:687). Furthermore, these differences in perspectives may have implications for students’ performance and their subsequent evaluations of both the course and the instructor (Goldstein & Benassi, 2006). An additional implication of the study may be that effective teaching also incorporates the ability of instructors to match the attributes that students identify as ideal for different types of instructional methods.

... differences in perspectives may have implications for students’ performance and their subsequent evaluations...

...effective teaching incorporates the ability of instructors to match the attributes that students identify as ideal for different types of instructional methods.

...(Brown, 2004)...examined the role of Neuro Linguistic Programming...in influencing students' perceptions of teaching effectiveness.

...“where students have different ... preferences, then even where other ... factors are met this leaves the student dissatisfied.”

...practical applications... focus on adopting [a variety of] teaching approaches that are more appropriate for their students' meta-programs, instructors are likely to enhance their teaching effectiveness...

... excellent teaching consists of a number of competencies related to content expertise, instructional design and delivery as well as interpersonal traits such as helpfulness, friendliness, open-mindedness, concern for students' learning and a genuine respect for students.

...while it is helpful to have a set of characteristics that can be used to identify excellent teaching “strict adherence to a set of principles...does not in itself establish effectiveness” it is more important to determine the degree to which an instructor's practice is guided by an explicit pedagogical framework...

To some extent, Goldstein and Benassi's study results reflect the findings of an earlier study (Brown, 2004) that examined the role of neuro-linguistic programming (NLP) preferences or meta-programs (i.e., ways of perceiving information and processing information, including the level of detail provided) in influencing students' perceptions of teaching effectiveness. In essence, the results of Brown's study suggest that “a teacher's meta-programs influence the approaches adopted in their teaching and these styles suit students with matching meta-program preferences” (2004:515). It follows that “where students have different ... preferences, then even where other ... factors are met [e.g., ability to communicate, teacher-student rapport], this leaves the student dissatisfied.” The practical applications of this research for evaluating effective teaching focus on supporting instructors to recognize that their meta-program preferences are likely to differ from those of at least some of their students; by adopting [a variety of] teaching approaches that are more appropriate for their students' meta-programs, instructors are likely to enhance their teaching effectiveness (Brown, 515).

In summary, excellent teaching consists of a number of competencies related to content expertise, instructional design and delivery as well as interpersonal traits such as helpfulness, friendliness, open-mindedness, concern for students' learning and a genuine respect for students. While there are some variations in the research results as to which characteristics are perceived by students and faculty as being strong predictors of effective or excellent teaching—including students' perceived value of the course—it also seems to be the case that effective teachers can compensate for deficiencies in one or two areas by demonstrating outstanding skills in other areas” (Young & Shaw, 1999:683). While it is helpful to have a set of characteristics that can be used to identify excellent teaching—and quantified to measure effectiveness—the Cornell Centre for Learning and Teaching cautions that “strict adherence to a set of principles [or behaviours or traits or skills] does not in itself establish effectiveness” (1997:40); rather it is more important to determine the degree to which an instructor's practice is guided by an explicit pedagogical framework and the frequency with which an instructor does or does not follow his or her pedagogical framework.

Developing Frameworks for Evaluating Excellence in Teaching

Researchers have noted the following weaknesses among existing frameworks: they are too...

In addition to identifying dimensions of excellent teaching, it is also important to develop a comprehensive, valid, and reliable framework for evaluating teaching effectiveness within a specific institution. Researchers have noted the following weaknesses among existing frameworks: they are too narrow in that they only focus on instructors' behaviours or actions, they do not pay attention to the person as a teacher (or for that matter, the teacher as a person), they often do not acknowledge the extent to which faculty construct and maintain a sense of professional identity through their disciplinary knowledge, they are not validated, do not take into account multiple perspectives of teaching effectiveness, and they are often not appropriate for contemporary approaches to teaching (Pratt, 1997; Young & Shaw, 1999; Braskamp, 2000; Tigelaar et al, 2004).

Guiding Principles for developing Evaluation Frameworks

In order to alleviate faculty mistrust and apprehension of the process—as well as to demonstrate that teaching activities are valued—it is of critical importance that evaluation frameworks are developed which are perceived as rigorous, credible, and respectful of the diversity of contexts and disciplines within colleges and universities...

As indicated in the introduction, faculty tend to respond with scepticism and mistrust when they perceive that they are being evaluated using too narrow a set of criteria, particularly in relation to summative evaluations. In order to alleviate faculty mistrust and apprehension of the process—as well as to demonstrate that teaching activities are valued—it is of critical importance that evaluation frameworks are developed which are perceived as rigorous, credible, and respectful of the diversity of contexts and disciplines within colleges and universities (Pratt, 1997). There appears to be some consensus in the field that evaluation frameworks (comprised of the policies, procedures and criteria) should be based on the following principles:

...evaluation frameworks should be based on the following principles...

- While evaluation should remain neutral with respect to individual style and beliefs, it must also acknowledge and respect diversity across faculty and disciplines.
- Evaluation should rest upon multiple and credible sources of data; the data should be able to (1) discriminate between individuals vis-à-vis specific competencies; (2) reliably and consistently measure specific competencies for individuals and between individuals, over time; and (3) provide information about the instructional context(s) in which a teacher functions the best.
- Evaluation should assess substantive aspects of teaching (i.e., disciplinary knowledge, underlying beliefs and intentions that shape decisions about course content) in addition to the technical (e.g., instructional design and delivery skills) component of teaching.

- Evaluation should contribute to improving teaching practice and career/professional development in general; to be effective, evaluation must be part of a larger faculty development and instructional improvement effort grounded in a philosophy that promotes continuous quality improvement.
- Evaluation should be done in consultation with the key individuals responsible for forwarding evaluative data and recommendations to the appropriate destinations within an institution.
- Evaluations should be based on clear communications of expectations by both the evaluator(s) and the instructor being evaluated.
- Evaluation of teaching should parallel other forms of judging scholarly work (Braskamp, 2000; Cornell Centre for Learning and Teaching, 1997; Pratt, 1997, Schaffner & MacKinnon, 2002).

...the specific criteria to be used in an evaluative framework—and how they are weighted— will depend, to some extent, on whether the evaluations are to be used for summative... formative purposes ...

At least three sources within the published literature contend that the specific criteria to be used in an evaluative framework—and how they are weighted— will depend, to some extent, on whether the evaluations are to be used for summative (i.e., for tenure or promotion purposes), or formative (i.e., to feedback for improving teaching practices) purposes (Braskamp, 2000; Cornell Centre for Learning and Teaching, 1997; Harrison, Douglas & Burdsal, 2004). The type of evaluation will also shape the question format on evaluation forms completed by students. Generally, summative instruments include mostly global items and use evaluative scales, whereas formative instruments focus more on specific items and are more likely to employ frequency scales (Cornell Centre for Learning and Teaching, 1997).

Opinions are mixed as to whether the two types of evaluations ought to be conducted separately or combined into one process.

Opinions are mixed as to whether the two types of evaluations ought to be conducted separately or combined into one process. One reason for keeping them separate is due to the apparently conflicting purposes of the two different types of evaluation: Summative evaluations are perceived as a public process intended to serve the purposes of administrators whereas formative evaluations are considered to serve individual teachers and are therefore confidential (Cornell Centre for Learning and Teaching, 1997). A second reason is that in cases where a consultant is brought in to oversee the evaluation process, combining the two types of assessment could result in a conflict of interest as to whose interests the consultant should be serving (Cornell Centre for Learning and Teaching, 1997). The disadvantages of treating the two evaluation functions separately are (1) the efficiency and cost of maintaining two separate processes, and (2) there is a danger that “the summative evaluation process may become too oriented toward comparing faculty with each other as a

The disadvantages of treating the two evaluation functions separately are (1) the efficiency and cost of maintaining two separate processes, and (2) a

danger that “the summative evaluation process may become too oriented toward comparing faculty with each other...at the expense of losing sight of individual instructors’ achievements in the improvement of teaching practices...”

...four major categories of an evaluation process...

Overall...consensus... on the basic elements to be included in the evaluation process... however, questions... specifics of which data sources should be included and how much weight is to be ascribed to them...

way of defining teaching effectiveness” at the expense of losing sight of individual instructors’ achievements in the improvement of teaching practices (Cornell Centre for Learning and Teaching, 1997:13). Pratt (1997) seems to imply that in the long run it may be more strategic—albeit more costly in terms of time and resources—to develop an elaborate evaluation framework that provides for rigorous and credible documentation from a variety of data sources.

The *Teaching Evaluation Handbook* developed by the Centre for Learning and Teaching at Cornell University (1997:16) has identified four major categories of an evaluation process:

- the range of characteristics of teaching that are evaluated (content expertise, instructional design, instructional delivery);
- the range of evaluation sources available to supply data (students and alumni, peers/colleagues, Dean, Chair, self);
- the criteria by which the data are to be evaluated (reliability, validity, objectivity of qualitative and quantitative data sources, weighting of various data and evaluation sources), and;
- the range of data types available for evaluation (e.g., teaching materials, student ratings, classroom observation, videotapes of teaching, structured interviews, teaching portfolios, reflective statements or memos by the instructor being evaluated).

Overall, there appears to be consensus in the literature on the basic elements to be included in the evaluation process as outlined above; however, questions as to the specifics of which data sources should be included and how much weight is to be ascribed to them, have produced varied study results and opinions in the existing literature. For example, some researchers (e.g., Pratt, 1997) contend that the evaluation process should be more substantive and include more input from the evaluation candidate regarding their beliefs and intentions. There are also some differences of opinion as to whether evaluation by one’s peers should play a greater role and carry more weight than evaluations from students, particularly in summative evaluations (Cohen and McKeachie, 1980, cited in Cornell Centre for Learning and Teaching, 1997; Nuhfer, 2003; Pratt, 1997).

Range of Competencies included in Evaluation Frameworks

Teaching is... a composite of content expertise, instructional design, and instructional delivery.

An array of characteristics within each category... should be assessed, using a variety of data types and evaluation sources...

...Cornell Centre for Learning and Teaching maintains that if teaching is to be acknowledged and rewarded as a valued activity, evaluation frameworks must also be able to determine whether instructors have, over a period of time, engaged in the following types of professional development activities...

Webbstock...specific criteria used in an evaluation framework... where a teaching portfolio is the main source of information, might be comprised of four core areas...

As indicated in the previous section, teaching is essentially a composite of content expertise, instructional design, and instructional delivery. An array of characteristics within each category—and particularly those characteristics that have been identified in the research as indicators of effective teaching—should be assessed, using a variety of data types and evaluation sources, in order to determine the quality of teaching. (Please see Appendix 1 for a list of criteria that might be used to evaluate each major category of instructional competencies. Note that these checklists of recommended questions/criteria have been drawn from various models discussed in this literature review.) Characteristics such as active participation in the academic community and evidence of life-long learning and scholarship have also been connected to overall teaching effectiveness (e.g., Braskamp, 2000; Hildebrand, Wilson & Dienst, 1971. cited in Cornell Centre for Teaching and Learning, 1997; Tigelaar et al, 1999) and could also be incorporated, to some extent, as criteria for assessing teaching excellence.

The Cornell Centre for Learning and Teaching also include criteria to assess improvement of practice over time, as part of their evaluation framework. The Cornell Centre for Learning and Teaching maintains that if teaching is to be acknowledged and rewarded as a valued activity, evaluation frameworks must also be able to determine whether instructors have, over a period of time, engaged in the following types of professional development activities: participating in teaching improvement opportunities, developing special teaching materials, revising courses, exploring alternative teaching methods, and/or seeking help in trying new teaching ideas (1997:41). Furthermore, it is suggested that factors such as teaching load as a percentage of the total job description should also be taken into consideration when evaluating how well an instructor has met departmentally based criteria such as whether an individual strives for excellence in teaching, has taken on responsibilities related to the teaching mission of the department or institution, and demonstrates initiative in recognizing and solving obstacles that hinder good teaching within the institution (Cornell Centre for Learning and Teaching, 1997).

Webbstock, writing from her experiences in a South African university, suggests that the specific criteria used in an evaluation framework, particularly where a teaching portfolio is the main source of information, might be comprised of four core areas:

- (1) A written statement of the instructor's philosophy of education;

- (2) Methods of teaching;
- (3) Methods of assessment;
- (4) Peer and student evaluation of the individual's teaching—and a choice of two of the following four "elective" areas for evaluation:
 - i) Evidence of ongoing professional development regarding teaching and learning;
 - ii) Evidence of development of new courses or extensive curriculum development;
 - iii) Evidence that the instructor shares teaching expertise with others; and ,
 - iv) Special recognition of an instructor's teaching expertise (1999: 173). According to the evaluation framework outlined by Webstock, each of these categories is assigned a rating from A (all criteria are clearly met) to E (criteria are not adequately met).

Evaluation Sources

Some research suggests that student evaluations conducted in the middle of the term are more useful for formative evaluations as it provides the instructor with feedback earlier in the term and an opportunity to make improvements in teaching practice before the semester ends.

Students are one of the main information providers of data on the quality and/or effectiveness of their instructors. In most post-secondary institutions, students are asked to complete evaluation forms near the end of a term or immediately after classes have finished—albeit usually before students write their final exams or receive their final grade for the course. Some research suggests that student evaluations conducted in the middle of the term are more useful for formative evaluations as it provides the instructor with feedback earlier in the term and an opportunity to make improvements in teaching practice before the semester ends. Former students and alumni might also be solicited for feedback on an instructor's effectiveness, although the research indicates it is best to ask former students to evaluate overall effectiveness and the value of the course than to ask them to recall the specifics of instructional delivery skills that occurred during the actual class (Cornell Centre for Learning & Teaching, 1997; Young & Shaw, 1999).

...while students are often asked for their impression of the instructor's knowledge of the subject area being taught, most researchers agree that assessments of content expertise ought to be conducted by peers in the discipline and colleagues in the department.

While students are often asked for their impression of the instructor's knowledge of the subject area being taught, most researchers agree that assessments of content expertise ought to be conducted by peers in the discipline and colleagues in the department. There is less overall agreement as to the role of students in evaluating a teacher's instructional design competencies; some researchers (e.g., Cornell Centre for Learning and Teaching, 1997; Marsh, 1987, cited in Young & Shaw, 1999) maintain that students are effective sources for evaluating certain aspects of course design skills, whereas other researchers contend that *all* aspects of instructional design competencies are best evaluated by peers and colleagues rather than students (Pratt, 1997). Pratt (1997) contends that most student evaluations of

instructors focus on the process of teaching rather than the outcomes of teaching—hence the data is more useful as feedback for improving teaching than providing insights into the effect of teaching. Other authors in the field seem to imply that information on both dimensions of evaluation (formative and summative) is helpful in certain contexts (Cornell Centre for Learning and Teaching, 1997).

... colleagues within the department and peers within an instructor's discipline are also important sources of information in the evaluation process. Braskamp... a faculty peer evaluation based on a "sitting beside" approach is more constructive in terms of building community - than a "standing over" approach to evaluations.

...faculty who are volunteered/recruited to play a role in peer evaluations need clear criteria and training if they are to provide useful, valid and reliable feedback...

Thus, colleagues within the department and peers within an instructor's discipline are also important sources of information in the evaluation process. Braskamp (2000) suggests that a faculty peer evaluation based on a "sitting beside" approach is more constructive—in terms of building community within the institution—and provides a slightly different perspective than a "standing over" approach to evaluations. Pratt also alludes to the need for a more collaborative and less competitive approach to peer evaluation (1997). As previously indicated, there is some agreement that faculty who are volunteered/recruited to play a role in peer evaluations need clear criteria and training—whether they are assessing written documentation or conducting classroom observations—if they are to provide useful, valid and reliable feedback; some researchers also suggest that the process would be more transparent if there is more open discussion between all parties involved pertaining to what constitutes teaching excellence (Webbstock, 1999) as well as identifying underlying assumptions and beliefs about teaching and learning that shape one's observations and interpretations of those observations (Pratt, 1997).

A third important source of information in the evaluation process is the instructor who is being evaluated.

Several researchers have reported that is advantageous for an instructor to supply a reflective document to assist evaluators.

A third important source of information in the evaluation process is the instructor who is being evaluated. Several authors maintain that since the "hard data by and of themselves cannot tell the complete story of an individual's teaching experience and development" (Cornell Centre for Teaching & Learning, 1997:17). Instructors are also considered to be the best source for documenting evidence of teaching effectiveness. Several researchers have reported that is advantageous for an instructor to supply a reflective document that:

- (1) will probe the core intentions and beliefs that inform decisions around the course planning decisions and documents;
- (2) will assist evaluators in the task of interpreting numerical student evaluation scores; and
- (3) provide information about an instructor's philosophy of teaching and learning (Cornell Centre for Learning & Teaching, 1997; Pratt, 1997; Webbstock, 1999).

Through discussions with colleagues and peer evaluators about the characteristics of excellent teaching, as well as

opportunities to have input into which classroom-based teaching competencies are evaluated or which course materials are selected for evaluation, the individual instructor takes a more pro-active role in their own evaluation.

Evaluation Data

...main categories of teaching competencies are captured in the following kinds of data: course materials, personal statements, student evaluations, feedback from peer reviews, and teaching improvement activities... student feedback from evaluation surveys have traditionally comprised the only type of data from students... researchers have suggested incorporating letters of recommendation from former students as acceptable forms of supporting data...

Data from peer reviews can include an assessment of written materials or observations of instructional delivery skills in the classroom.

A major source of evaluation data submitted by the instructor is a teaching portfolio...

The purpose of a teaching portfolio is to "enable faculty members to display their teaching accomplishments for examination by others"... the portfolio should represent enough of the key dimensions of teaching to highlight teaching as a scholarly activity...

According to the literature on evaluation frameworks, the main categories of teaching competencies are captured in the following kinds of data: course materials, personal statements, student evaluations, feedback from peer reviews, and teaching improvement activities (Cornell Centre for Learning & Teaching, 1997; Pratt 1997; Webbstock, 1999). While student feedback from evaluation surveys have traditionally comprised the only type of data from students, some researchers have suggested incorporating letters of recommendation from former students as acceptable forms of supporting data; however, effective use of qualitative data requires more preparation to put the information into a format that is useful, concise, and protects the anonymity of the student (Cornell Centre for Learning and Teaching, 1997: 34).

Data from peer reviews can include an assessment of written materials (e.g., course materials used, texts/readings selected, samples of assignments, exams, grading schemes, etc) or observations of instructional delivery skills in the classroom. While some practitioners in the evaluation field recommend orienting observational criteria toward the technical skills related to planning, time-management, instructor-group interactions, and presentation of material, others are of the view that the focus should be on how instructional delivery skills are carried out and the extent to which the teaching behaviours agree with the underlying intentions and beliefs (Cornell Centre for Learning and Teaching, 1997; Pratt, 1997).

A major source of evaluation data submitted by the instructor is a teaching portfolio which might include a reflective statement and work samples as well as evidence of an on-going commitment to improving teaching practice (e.g., documentation of new or innovative teaching approaches, reflective writing, samples of curriculum development, documentation showing participation in professional development activities, etc.) The purpose of a teaching portfolio is to "enable faculty members to display their teaching accomplishments for examination by others" (Cornell Centre for Learning and Teaching, 1997:20); furthermore the portfolio should represent enough of the key dimensions of teaching to highlight teaching as a scholarly activity (Cornell Centre for Teaching & Learning, 1997; Pratt, 1997; Webbstock, 1999). A

widespread observation in the literature is the need to balance inclusiveness with brevity when compiling teaching portfolios in order to avoid overwhelming evaluators. Centres of teaching excellence in many Canadian (e.g., University of Manitoba, University of Alberta) and America provide support—in the form of published guidelines or workshops—for faculty who wish to create a teaching portfolio.

Reliability and Validity of Evaluation Criteria

...criteria must be both reliable and valid for evaluation frameworks to be deemed credible.

...regardless of the data being reviewed, establishing a replicable protocol and standardizing rating procedures and criteria are crucial for ensuring the validity, fairness and accountability of the evaluation process...

...where qualitative data... are used for summative evaluation purposes, the validity and reliability of such data are improved by providing a structured set of questions...

There is general consensus in the literature that criteria must be both reliable and valid for evaluation frameworks to be deemed credible. Questions have been raised—and debated—with regard to the reliability and validity of both student and peer evaluation data. For example, a study by Centra (1975; cited in Cornell Centre for Learning and Teaching, 1997) found that if peer observers conducting classroom visits were untrained in their task, their observations were statistically less reliable. Subsequent studies have shown that training evaluation sources—whether students or peers—prior to undertaking assessment related tasks seems to reduce rater errors and enhance reliability and accuracy (Young & Shaw, 1999). The Cornell Centre for Learning and Teaching notes that, regardless of the data being reviewed, establishing a replicable protocol and standardizing rating procedures and criteria are crucial for ensuring the validity, fairness and accountability of the evaluation process (1997:36). To this end, a number of researchers have developed clear guidelines for processes such as the administration of student evaluation questionnaires, the reporting of summative evaluation scores (Cornell Centre for Learning and Teaching, 1997), and classroom observations conducted by peers or colleagues (Cornell Centre for Learning and Teaching, 1997; American Association of Higher Education, 1995 cited in Pratt, 1997). It has also been suggested that where qualitative data (e.g., letters from former students or alumni) are used for summative evaluation purposes, the validity and reliability of such data are improved by providing a structured set of questions with regard to the teaching competencies (or value of the course) that are to be evaluated (Cornell Centre for Learning & Teaching, 1997).

Transparency of Evaluation Procedures: Some Examples

...in recent years there has been a trend toward making evaluation policies and procedures more available and accessible on universities' websites...

In the past, evaluation policies and procedures were embedded in personnel policies and faculty agreements; however, in recent years there has been a trend toward making evaluation policies and procedures more available and accessible on universities' web sites both in Canada and America. The move toward creating—and publicizing—comprehensive evaluation frameworks has not developed evenly across institutions within higher education. Some examples of institutions with comprehensive, highly structured evaluation frameworks are outlined below.

Arizona State University

Teaching evaluations at this university are comprised of activity reports, student evaluations, portfolios and written goal-based performance agreements which are evaluated by a personnel committee. General criteria for evaluating teaching are determined by each school, however there are specific expectations, such as teaching on a regular basis and staying current within one's discipline, that are institution-wide. Faculty are evaluated on a scale ranging from unsatisfactory to exemplary in three core areas—teaching, research and service—as well as overall performance. Guidelines have been established to outline how performance is to be determined, as well as identifying the consequences or implications of receiving an unsatisfactory evaluation in one or more of the three core areas. These performance level evaluations underpin a merit based system of discretionary salary allocations. (For more details, see <http://www.asu.edu/aad/manuals/acd/acd506-10.html>.)

Cornell University

The evaluation framework used at this institution was developed by the Cornell Centre for Learning and Teaching. The framework is described in a handbook ('Teaching Evaluation Handbook') that is maintained on the Centre's web site. The purpose of the handbook is to ensure uniformity of practice throughout the institution, as well as to help faculty who are seeking tenure to document their teaching activities, support tenure committees, and to ensure that teaching evaluations are fair, rigorous, and comprehensive. See www.clt.cornell.edu for additional information.

University of Toronto, Faculty of Arts & Science

Teaching evaluations at this institution use multiple types and sources of data: evaluations by current students, structured feedback from prior students, peer evaluations of classroom performance, and confidential assessments from co-teaching colleagues. The institution implies that evaluations are conducted within a framework supporting faculty to become more effective teachers. Teaching deficiencies that come to light in the evaluation are identified by the instructor, in consultation with the department chair so that resources can be provided to improve the instructor's teaching effectiveness. (See <http://www.teaching.artsci.utoronto> for a more complete description of the framework.)

University of Saskatchewan

Teaching evaluations at the University of Saskatchewan use multiple types and sources of data. The evaluation process and the data to be included are clearly spelled out in their collegial processes policies, and are posted on the web site for the Office of the Provost & Vice-President Academic. (See <http://www.usask.ca/vpacademic/processes.shtml>)

University of Alberta

Instructors at the University of Alberta are evaluated in terms of their performance in three areas: teaching, research, and service. While the criteria for evaluation are contained within the Faculty Agreement document, the general principles and process is outlined in their GFC Policy Manual and available online. Faculty evaluations are multi-faceted; they are comprised of student ratings captured in a standardized form and process as well as data from one or more of the following sources of data: administrators, peers, self, undergraduate and graduate students, and alumni. The university's Teaching Services recommends—and provides support for—developing a teaching portfolio (dossier) as the basis for providing information in the category of "input from self". (For more information, refer to their online resources at <http://www.uofaweb.ualberta.ca/gfcpolicymanual/content>)

Facilitating Teaching Excellence

...teaching evaluation processes that are not grounded in a philosophy of promoting teaching excellence or do not provide the professional development opportunities to support improving teaching practice are:
(1) viewed with scepticism; and,
(2) tend to generate anxiety in response to the evaluation process and its outcomes.

As indicated in an earlier section of this literature review, teaching evaluation processes that are not grounded in a philosophy of promoting teaching excellence or do not provide the professional development opportunities to support improving teaching practice are:

- (1) viewed with scepticism and
- (2) tend to generate anxiety in response to the evaluation process and its outcomes.

Following are examples of universities (or colleges) in which the stated goals of encouraging and fostering excellent teaching are reflected in the resources that have been created to support faculty in becoming excellent teachers.

University of Manitoba, University Teaching Services

This service unit provides numerous courses designed to assist faculty to develop their teaching skills, develop teaching portfolios, identify learning and teaching preferences (e.g., meta-programmes), and gain a better understanding of specific student groups (e.g., international programs, first year students). The University Teaching Service also provides faculty with information on the tenure application process.

University of Saskatchewan, Gwenna Moss Centre for Teaching Effectiveness

This centre provides instructors with print and online resources on excellent teaching practices, teaching and learning guides, guidelines for developing teaching portfolios, and orientation sessions for new instructors and academic staff.

University of Alberta, University Teaching Services

This service unit is recognized and supported by the institution for its role in developing and improving teaching and learning practices and enhancing research in university teaching. The service unit provides a number of services to support instructional design and delivery skills, developing teaching portfolios, mentoring and peer consultation processes, orientations and workshops, and a resource centre. (For additional information refer to their web site at <http://www.ualberta.ca/~uts/Resources>)

Cornell University, Centre for Learning and Teaching

This centre has developed programs to support both learners and teachers. Services specifically for faculty instructional support include seminars for improving planning, classroom presentation skills and interaction skills, mid-semester student evaluations to provide feedback on teaching quality and effectiveness, and opportunities to have classes videotaped for self – evaluation purposes.

University of Colorado at Boulder, Faculty Teaching Excellence Program

This program offers similar types of programs to those offered by the Cornell Centre for Learning and Teaching. The program appears to offer seminars on the practical aspects of teaching, provide resources on teaching related to specific programs or departments. Services such as videotaping a class session, classroom observation, and administering surveys are available to provide instructors with an opportunity to evaluate their own teaching performance.

References

Braskamp, Lawrence A. (2000) Toward a More Holistic Approach to Assessing Faculty as Teachers. *New Directions For Teaching and Learning* 83:19-33.

Brown, Nigel (2004) What makes a good educator? The relevance of meta programmes. *Assessment & Evaluation in Higher Education* 29(5):515-533.

Cornell Centre for Teaching and Learning (1997) *Teaching Evaluation Handbook (3rd ed.)* Ithaca, N.Y.: Office of Instructional Support, Cornell University.

Feldman, K.A. (1988) Effective College Teaching from the students' and faculty's view: Matched or mismatched priorities? *Research in Higher Education* 28(4):291-329.

Goldstein, Gary S. and Victor A. Benassi (2006) Students' and Instructors' Beliefs about Excellent Lecturers and Discussion Leaders. *Research in Higher Education* 47(6):685-707.

Harrison, Paul D., Deanne K. Douglas, and Charles A. Burdsal (2004) The relative merits of different types of overall evaluations of teaching effectiveness. *Research in Higher Education* 45(3):311-323.

Nehari, Miriam and Hilary Bender (1978) Meaningfulness of a Learning Experience: A measure for educational outcomes in higher education. *Higher Education* 7:1-11.

Nuhfer, Edward B. (2003) *Of what value are student evaluations?* Pocatello, Idaho. Center for Teaching and Learning, Idaho State University.

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

Schaffner, Monika and Fiona J.D. MacKinnon (2002) *A Standards-Driven Approach to Faculty Evaluation: The Conflict of Change*. (Paper presented at the Annual Meeting of the American Educational Research Association.) New Orleans, Louisiana.

Tigelaar, Dineke E.H., Diana H.J.M. Dolmans, Ineke H.A.P. Wolfhagen and Cees P.M. Van der Vleuten (2004) The development and validation of a framework for teaching competencies in higher education. *Higher Education* 48:253-268.

Webbstock, Denyse (1999) An Evaluative Look at the Model used in the Assessment of Teaching Quality at the University of Natal, South Africa: reflections, rewards and reconsiderations. *Assessment & Evaluation in Higher Education* 24(2):157-179.

Young, Suzanne and Dale G. Shaw (1999) Profiles of Effective College and University Teachers. *The Journal of Higher Education* 70(6):670-686.

This page intentionally left blank.

Appendix 1: Criteria to Assess Categories of Teaching Competencies

Component	Content Expertise	Instructional Design	Instructional Delivery
What is being assessed?	Mastery of content/discipline Depth & breadth of knowledge Ability to analyze & synthesize information Appropriateness of course goals/objectives Selection of course content (appropriate for level taught) Engagement with subject (enthusiasm)	Organization and planning skills Course design and development Skills for evaluating student learning	Presentation skills—delivery of information to students Classroom interactions with students Interactions with individual students Presents information in variety of ways
Possible criteria for evaluation*	Contrasts implications of theories Discusses points of view other than their own Gives additional references for more interesting, complex points Discusses recent developments in field Presents origins of ideas/concepts Presents facts, concepts, related fields Emphasizes conceptual understanding Can suggest reading in any area of general field Well read beyond subject taught Know about developments in other fields Sought by colleagues for advice on academic matters Course objectives are congruent with department curricula Course objectives are clearly stated Reading list is current and includes works of recognized authorities	Course objectives are clearly identified Syllabus adequately outlines sequence of topics covered Course requirements, attendance policies & grading standards stated clearly in syllabus Outline and topic sequence are logical Written assignments/projects reflect course goals Exam content is representative of course content and objectives Exam questions are clearly written; tests are well designed and carefully selected Assignments & tests are graded fairly & returned promptly Grade distribution pattern is appropriate for course level	Explains information clearly—including directions for assignments, class activities Give lectures that are easy to take notes in (or provides class outlines for students) Summarizes major points, draws attention to important points Encourages class discussions, sharing of students' experiences Know whether class is understanding material or is bored/confused Has a genuine interest in students; is friendly, accessible/approachable, respects students Encourages students to think about the subject, intellectually challenges students Uses class time efficiently and ties things together Creates a comfortable learning environment Uses variety of props, audiovisual equipment, etc.
Types of Data Used	Reading list, syllabus, , reflective memo, peer observations / evaluations, student evaluations	Samples of course exams/ assignments, lesson plans, syllabus, course objectives, peer evaluations	Student evaluation forms Classroom observation by peers
Data Sources	Instructor, peers, students, Dept Chair	Colleagues and/or peers in field, students	Students, Peers/Colleagues

*Note: this is not an exhaustive list of criteria.

Sources: Cornell Centre for Learning and Teaching (1997), Hildebrand, Wilson & Dienst (1971, cited in Cornell Centre for Learning and Teaching, 1997), Pratt (1997).