

BACKGROUND: METHODS FOR THE ASSESSMENT OF TEACHING

INTRODUCTION

- The Complementary Teaching Assessment Project Team (CTAPT) conducted a literature review and environmental scan of U15 current and best practices for the assessment of teaching, other than student evaluation of teaching (SET).
- CTAPT proposes that the University of Waterloo adopts Teaching Dossiers (TD) and Peer Review of Teaching (PRT) as complementary methods, for they provide valuable and robust ways for documenting evidence of teaching. CTAPT recognizes that some units at the UWaterloo already employ PRT or TD.
- This Backgrounder outlines CTAPT's key findings and provides an overview of current practices and a framework for implementing and supporting complementary methods for the assessment of teaching. It also includes "Fact Sheets" that describe best practices for Teaching Dossiers and Peer Review, and how to use them to assess teaching effectiveness.
- For more details on CTAPT's research findings, please consult the full report here: *CTAPT Committee Report on Assessment Methods*.

CORE ELEMENTS OF AN EFFECTIVE EVALUATION FRAMEWORK

- The assessment of teaching is done for two main purposes: i) providing feedback for faculty growth, teaching development or improvement, and ii) providing evaluative information for personnel decisions (Arreola 1995: 2; Wright et al. 2014: 36).
- Much of the literature advocates for an "improvement" or "growth-orientated" evaluation framework (e.g. Arreola 2007, 1995; Chism 1999; Seldin 2010, 2006, 1984; Wright et al. 2014).
- A recent and thorough study by Wright et al. (2014) identifies some reoccurring themes in effective evaluation frameworks that align with the mission of CTAPT and the vision at the University of Waterloo:
 1. **A MULTI-FACETED OR COMPLEMENTARY APPROACH:**
 - Utilizing multiple sources of evidence - from students, peers, and instructor - and **multiple methods** such as surveys, peer observations, and teaching dossiers increases reliability and fairness (see Berk 2014: 88; Arreola 2007; Chism 1999, 2007; Hubball and Clark 2011; Seldin 1999, 2007).
 2. **A SHARED UNDERSTANDING OF TEACHING EFFECTIVENESS:**
 - This definition should be "contextual, evolving, and periodically reviewed" (Devlin and Samarawickrema 2010), consider "faculty values" and include evaluation criteria (Arreola 1995: 1; Berk 2006; Wright et al. 2014: 14).
 - The literature acknowledges the fact that the notion of 'good' or 'effective' teaching is ambiguous in many teaching evaluation programs.
 - CTAPT developed a **definition of teaching effectiveness** based on research conducted and consultation with the UWaterloo community. Results from the consultation survey and finalized definition are available here: [CTAPT Teaching Effectiveness Survey Results – Campus Report](#).
 3. **ROBUST FEEDBACK CYCLES AND SUPPORT**
 - A teaching evaluation process with "...robust feedback cycles that are integrated into evaluation and instructional improvement programs" where faculty development programmes support feedback and improvement cycles (Wright et. al. 2014: 5, 17).
 4. **MULTI-LEVEL LEADERSHIP AND ENGAGEMENT:**
 - Multi-level leadership in fostering a culture that values and rewards teaching, as well as consultation and communication are key to an effective integrated and multifaceted approach (Wright et al. 2014: 17, 18).

CURRENT AND RECOMMENDED PRACTICES FOR THE ASSESSMENT OF TEACHING

- Although Berk (2005) reviews 12 potential sources of evidence to assess teaching effectiveness, research confirmed most higher education institutions mainly use one or more of the following three: student evaluations (SET), self-evaluation such as teaching dossiers (TD), and peer review of teaching (PRT).

- The teaching dossier is the most commonly used complementary method in Canada for summative evaluation while peer review is more commonly used for formative purposes. A mixed method study that included responses from participants at 16 Universities in Ontario revealed the following (Wright et al. 2014: 40, 42; Gravestock 2001):
 - **Summative:** 82% use SETs, **50% use teaching dossiers**, 29% use other self-evaluation instruments, 20% use peer observation.
 - **Formative:** **67% use peer observation**, 43% use self-evaluation instruments, and 31% use teaching dossiers.
- Effective and well-supported use of PRT or TD leads to positive outcomes and the following benefits: enhances a scholarly approach to teaching, reflective practice and professional development; leads to innovations or changes to teaching practices; facilitates opportunities for dialogue and collegiality; strengthens the validity and reliability of teaching evaluation through triangulation of evidence.¹
- The main concerns with PRT and/or TD identified in the literature include time commitments, quality of feedback, and the lack of clear standards, criteria, and tools, which relate to additional concerns about validity, subjectivity, and bias.²
- PRT is a valuable complementary method for providing evidence on dimensions of teaching effectiveness students are unable to assess.
- A TD is an ideal complementary method to use, as components or sections of a TD align with the four dimensions of teaching effectiveness developed by CTAPT. It also provides a method for compiling and contextualizing multiple sources of evidence (i.e. from self, peers, students, and the literature) and a framework for reflecting on teaching practices holistically.
- Brookfield (1995) argues that critical reflection using the four lenses (self, peers, students, theoretical literature) is what distinguishes excellent teachers. Hubball and Clarke (2011: 1) also argue for the value of using a **scholarly approach** to teaching in research-intensive universities, which involves consulting the literature and peers to determine and implement best practices and disciplinary approaches, then reflecting on and assessing those practices (differentiated from SoTL).

BEST PRACTICES

- CTAPT has developed **PRT and TD Fact Sheets**, which summarize the *key components*, benefits and potential concerns, and *best practices for maximizing benefits and addressing concerns* (see page 3-6).
- The Fact Sheets are based on findings from the literature review, including evidence from studies reporting on the development, implementation, and evaluation of pilot programs (particularly of PRT), which offer insight on how they addressed concerns and overcame implementation challenges.
- They also draw on several U15 evidence-based pilot study reports and/or guidelines (e.g. Kenny et al. 2018; Richard 2018; University of Toronto 2017a, 2017b).
- CTAPT has mapped the three methods of assessment (PRT, TD, SET) to the Dimensions of Teaching Effectiveness to show where each method should be used (see **Fact Sheet, page 7**). The tables demonstrate how utilizing Peer Review of Teaching and Teaching Dossiers along with student surveys enables instructors to highlight evidence of teaching effectiveness more comprehensively and contextualize results, leading to a more valid and reliable representation of teaching.

NEXT STEPS: PHASE 2 OF CONSULTATIONS

- In October 2019, CTAPT will be inviting faculty members to participate in a consultation through either an in-person session or online survey. The purpose of this consultation is to gather input from faculty members on what they would need to implement and support Peer Review of Teaching and Teaching Dossiers in terms of both using these methods and evaluating evidence from these methods.
- Check your email for a call to participate, and please review the Fact Sheets below before participating – Thank you!

¹ See Barnard 2001; Bell and Cooper 2013; Chism 2007: 6; Gormally et al 2014: 188; Iqbal 2014: 113-5; Mager et al 2014 Schonwetter et al. 2002: 91; Seldin 2010: 43; Smith 2014; Teoh et al. 2016: 1; Thomas et al. 2014: 150.

² See Barnard 2001; Bell and Cooper 2013; Chism 2007: 6; Gormally et al 2014: 188; Iqbal 2014: 113-5; Mager et al 2014 Schonwetter et al. 2002: 91; Seldin 2010: 43; Smith 2014; Teoh et al. 2016: 1; Thomas et al. 2014: 150.

TEACHING DOSSIER FACT SHEET

WHAT IS A TEACHING DOSSIER?

- A teaching dossier (aka teaching portfolio) is to teaching what a CV is to research. It describes documents and materials, which collectively suggest the scope, quality, and impact of teaching (Seldin 2010).
- Presents an integrated summary of one's teaching philosophy, approaches, accomplishments and effectiveness through a reflective narrative and curated assembly of "robust and accurate evidence" based on multiple sources (Kenny et al 2018: 6). Key characteristics of a dossier are:
 - A central narrative piece, which describes your teaching philosophy and highlights important components and evidence in the dossier.
 - A purposeful representation of 'best work' in Appendices illustrating your performance and effectiveness.
- Varies in length anywhere from 5-20 pages depending on stage of career and purpose, but typically includes the following components and range from 5-8 pages (based on Holmes, Focused Dossier Proposal):
 1. **Teaching Philosophy** (central narrative) "What I believe and why" e.g. beliefs about teaching, approach, goals, strategies, methods.
 2. **Teaching Responsibilities** (strategies and actions) "What I do and how" e.g. list of courses taught, supervisory roles, curriculum development, teaching methods / practices and activities.
 3. **Results and Accomplishments** (impact or outcomes) "What were the results and what did I learn? e.g. teaching awards, publications, contributions; changes or innovations to teaching, and reflections on methods or changes.
 4. **Professional Development and Future Plans** "How will I improve and grow?" e.g. sessions or conferences, leading teaching groups, workshops, SoTL, publications.
 5. **Appendices of Evidence** (artifacts) "What evidence best demonstrates my strategies and accomplishments?" e.g. student surveys, awards, letters/comments, syllabi, assignments, PRT. Includes evidence from Self, Peers, and Students
- Teaching dossiers are evaluated holistically, as criteria may be found in multiple sections; evaluation should be based on pre-determined criteria and evidence provided in the dossier to ensure reliability.
- Resources might include Dossier Guidelines and evaluation tools such as guiding questions, checklists, or rubric for the different components.

WHY USE A TEACHING DOSSIER?

- Best tool for a multifaceted or complementary approach, it is the most effective way to integrate multiple forms and sources of evidence, providing a more accurate and reliable base for formative and summative decisions than using only one source, such as student evaluations (Berk 48, 49).
- Provides a method for supporting claims empirically with evidence and for contextualizing multiple forms of evidence.
- Facilitates a systematic, scholarly approach to teaching and learning – it offers both a method for collecting and documenting evidence and a framework for reflecting on teaching in order to innovate, grow, or improve accordingly.
- Instructor composed, a teaching dossier balances systematic evaluation with faculty control.
- Teaching dossier components align with teaching effectiveness dimensions (DEED) proposed by CTAPT.

POTENTIAL CONCERNS OR BARRIERS TO BE OVERCOME BY THE UNIVERSITY OF WATERLOO

- Time and effort to develop and review.
- Unfamiliarity with reflexive writing process.
- Need for guidelines or tools for evaluating a teaching dossier.
- Concerns about validity, reliability, subjectivity, and bias towards certain writing styles or evaluator's own / particular pedagogies and teaching methods or approaches.

BEST PRACTICES

- Clear procedures and process that acknowledge and address time.
- Well-defined and known criteria for teaching and learning.
- Gather evidence early and often.
- Providing evidence of teaching combined with clear procedures for evaluating evidence ensures reliability and reduce bias.
- Available resources such as guidelines for developing and tools for evaluating teaching dossiers.
- Training and support for developing AND evaluating teaching dossiers.
- Awareness of pedagogical and disciplinary practices.
- A culture that values and rewards teaching.

PEER REVIEW OF TEACHING FACT SHEET

WHAT IS PEER REVIEW OF TEACHING (PRT)?

- “Peer review of teaching is a systematic, reflective process through which teaching colleagues offer instructors feedback about their teaching for either formative or summative purposes, based on multiple forms of data” (Chism 2017 in Wright 2014: 21).
- We use PRT specifically to mean peer observation of teaching, which is an intentional observation process where a peer observes an instructor with the aim of providing critical feedback, based on predetermined goals and purpose (Thomas et al. 2014: 117).
- PRT review is used for different purposes or outcomes; PRT feedback may be used for development or improvement purposes (formative feedback) or for evaluation and appraisal (summative feedback) (see Gosling 2002).
- Peer review generally involves three steps or stages:
 1. **Pre-observation meeting** (setting of observation goals and expectations, logistics, etc.; observer review of relevant materials such as syllabi, assignments, etc.).
 2. **Observation** (in-class or online utilizing pre-determined tool).
 3. **Post-observation meeting** (instructor self-reflection, observer’s constructive feedback, discussion of strengths and opportunities as ‘critical friends,’ written report when applicable).
- Observations may be done by a peer, in reciprocal pairs, as triads (includes external), or draw from peer review teams.
- Observers utilize tools or instruments such as checklists (e.g. [CTAPT Definition of Teaching Effectiveness](#)), rating scales or rubric, templates, general or focused guiding questions, or narrative logs to document evidence from observations.
- Frequency of peer review varies and depends on context; however, it should occur within a known cycle and ideally consist of multiple observations and/or observers.

WHY USE PEER REVIEW OF TEACHING?

- Facilitates professional development and growth:
 - Shown to enhance engagement in reflective practices.
 - Shown to facilitate changes or innovation to teaching practices.
- Strengthens validity and reliability of the assessment of teaching by providing important evidence students are not able or knowledgeable enough to assess; provides “triangulation” of data.
- Enhances opportunities for dialogue about teaching and learning and facilitates collegiality.
- Is often mutually beneficial for both the instructor being observed and the observer.

POTENTIAL CONCERNS OR BARRIERS TO BE OVERCOME BY THE UNIVERSITY OF WATERLOO

- Time commitment – how to balance effectiveness of method with resources and scheduling availability.
- Unsure of procedures and/or how PRT will be used in a high-stake context.
- Lack of clear guidelines, criteria, tools or instruments, and training.
- Feedback that is too vague or too positive and thus not useful.
- Concerns about validity, reliability, and subjectivity.
- Reviewer bias towards own style or pedagogical approaches to teaching (or are not knowledgeable of certain pedagogical approaches).
- Concerns about observer’s lack of evaluation/observation skill or experience.

BEST PRACTICES

- Establish clear processes and practical procedures (in consultation with Faculties) that acknowledge and value time for peer review.
- The purpose, goals, and use (formative or summative) of peer review are clear.
- Incorporate opportunities for formative peer review of teaching and multiple observations and observers.
- Well-defined and known criteria, within context (e.g. what is effective teaching effectiveness? what approaches are used within the discipline?)
- Guidelines and tools are available and support quality feedback that is timely, relevant, descriptive, constructive, sensitive, and confidential for formative.
- Training for conducting observations, using tools, and providing feedback is available and encouraged.
- Institutional leadership - a culture that values and rewards teaching.
- Peer review is one component of a multifaceted and scholarly approach to the assessment of teaching.

USING COMPLEMENTARY METHODS FOR THE ASSESSMENT OF TEACHING

- No one source of evidence is the best source of information for ALL teaching activities.
- Utilizing multiple methods and sources of evidence increases the reliability, validity and fairness of the assessment of teaching; when possible, triangulate multiple sources or forms of evidence.
- Using Peer Review of Teaching (PRT) and Teaching Dossiers (TD) along with data from student course perception surveys (SCPS) enables instructors to document evidence of teaching effectiveness more comprehensively and provide context for results (see Table 1) .

TABLE 1: OVERVIEW OF A COMPLEMENTARY APPROACH TO ASSESSMENT

Dimension of Teaching Effectiveness	Sources	Methods for Providing Evidence / Information
DESIGN	Instructor Peer	TD: Instructor provided descriptive data explaining beliefs and rationale for teaching approach (teaching philosophy) and methods used with evidence, such as syllabi, assignments, or learning activity examples. PRT: Peer review of course materials, such as syllabi and learning materials or activities
EXECUTION	Peer Instructor Student	PRT: Peer observation of the implementation of teaching methods and approaches, review of materials, pre-and post-meeting with instructor TD: Instructor descriptions on using methods, strategies and approaches, and reflection on implementation. SCPS: Student's experience with specific aspects using appropriate student survey questions
STUDENT EXPERIENCE	Student Peer Instructor	SCPS: Appropriate student survey questions about student's perceptions of their own experience PRT: Peer observation based on set goals for observation and using appropriate tools, review of course materials TD: Instructor narrative on strategies and approaches used and reflection on implementation; contextualization or responses to survey results (and PRT if completed)
PROFESSIONAL DEVELOPMENT	Instructor Chair Peer	TD: Description of scholarly approach - reflection and application of practices; evidence of professional development (e.g. teaching workshops, conferences, SoTL publications, teaching awards, mentorship, etc.) Chair: Letters of support, knowledge of instructor's profession development, involvement in committees, curriculum work, or mentorship Peers: Letters of support in relation to mentorship, contributions, etc.

TABLE 2: MAPPING METHODS TO DIMENSIONS OF TEACHING EFFECTIVENESS

- According to Arreola's "rule of thumb" for selecting sources of evidence, "**...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**" (Arreola 2007: 47 in Murphy et al. 2009: 228; see also Chism 1999; Mills in Seldin 1999; and Seldin 2006).
- Keeping in mind that more than one method may be a source of evidence, the Table below outlines which method provides the **most direct source of evidence** for each descriptor of teaching effectiveness and is thus a "primary source" of evidence. Primary sources are not more important, rather, they provide the *most direct* source of evidence; secondary sources are those that provide additional observable evidence.
 - Primary Sources are indicated with two checkmarks (✓✓)
 - Secondary Sources are indicated with one check mark (✓)
 - Not a source is indicated with an "X"

DIMENSION OF TEACHING EFFECTIVENESS	PEER REVIEW	TEACHING DOSSIERS	STUDENT COURSE PERCEPTION SURVEY
DESIGN			
<i>Planning</i>			
Builds course around a design framework from the literature, from the field of study, or created by the course author	✓	✓✓	X
Clearly defines course-level and activity-level learning objectives	✓	✓✓	✓ ⁱ
Includes learning material that reflects current scholarship from the field or that is clearly relevant	✓	✓✓	X
Structures material in a logical and coherent order	✓	✓✓	✓*
Sets pacing, workload and performance standards appropriate for the course level and topic	✓	✓✓	✓✓ ⁱⁱ
Includes experiential components, professional connections, or practical applications, when possible.	✓	✓✓	X
Plans a variety of teaching/learning strategies to promote student engagement and deep approaches to learning	✓	✓✓	X
Incorporates a diversity of experiences, viewpoints, and backgrounds in course materials	✓	✓✓	X
Adheres learning materials, activities, and assessments to University accessibility policies	✓	✓✓	X
<i>Alignment</i>			
Aligns course design with program expectations	✓	✓✓	X
Aligns course objectives and learning outcomes with course content and delivery	✓	✓✓	X
Develops fair and equitable assessment methods that align with course objectives and outcomes	✓	✓✓	✓✓ ⁱⁱⁱ

NB: Piloted SCPS questions that correlate with descriptors are indicated with a footnote (see also Appendix). An asterisk (*) indicates that there is no corresponding pilot questions at this time, but where surveys or other student evidence may be a sources of evidence.

- i. "The instructor identified the intended learning outcomes for this course" (Based on Student Course Perception Survey questions piloted by [CEPT2](#), accessed September 19, 2019).
- ii. "The course workload demands were..." [Select a choice from a scale of very low to very high] (ibid).
- iii. "The intended learning outcomes were assessed through the graded work" (ibid).

DIMENSION OF TEACHING EFFECTIVENESS	PEER REVIEW	TEACHING DOSSIERS	STUDENT SURVEYS
EXECUTION			
<i>Communication</i>			
Communicates course-level and activity-level objectives as well as teaching/learning approach and rationale to students	✓	✓✓	✓✓ ⁱ
Describes and explains material clearly using a pace appropriate to the context	✓✓	✓	✓✓ ^{iv}
Demonstrates enthusiasm for the subject	✓✓	✓	✓✓*
Uses technology, media or other teaching tools effectively	✓✓	✓	✓*
<i>Student Engagement</i>			
Promotes student participation, peer interactions, or other active engagement with course content	✓✓	✓✓	✓*
Uses teaching/learning strategies that encourage student engagement and deep approaches to learning	✓✓	✓✓	✓*
<i>Variety of Elements</i>			
Adapts to evolving classroom contexts	✓	✓✓	X
Adopts a variety of instructional practices, content types, and assessments that recognize diversity of learners	✓	✓✓	X
<i>Assessments and Feedback</i>			
Enables students to prepare for assessments through instructional practices	✓	✓✓	✓✓ ^v
Communicates clear expectations and instructions for assessments	✓✓	✓	✓✓*
Provides performance feedback in a timely manner	X	✓	✓✓ ^{vi}
Provides directions for student improvement individually or collectively	X	✓✓	✓✓*

iv. "The instructor helped me to understand the course concepts" (ibid).

v. "The course activities prepared me for the graded work" (ibid).

vi. "Graded work was returned in a reasonable amount of time" (ibid).

DIMENSION OF TEACHING EFFECTIVENESS	PEER REVIEW	TEACHING DOSSIERS	STUDENT SURVEYS
STUDENT EXPERIENCE			
<i>Rapport</i>			
Fosters a supportive learning environment	✓✓	✓	✓✓ ^{vii}
Establishes a climate of intellectual openness	✓✓	✓	✓✓ ^{vii}
Shows concern for students' success and wellbeing	✓✓	✓	✓✓
Interacts professionally and respectfully with students	✓✓	✓	✓✓
<i>Responsiveness</i>			
Provides sufficient opportunities for student contact inside and outside of class	✓	✓✓	✓✓
Responds to student inquiries and questions in an appropriate timeframe	X	✓✓	✓✓
<i>Diversity</i>			
Promotes the success of all students by addressing the variety of their experiences, viewpoints, and backgrounds	✓✓	✓✓	✓
<i>Engagement and Learning</i>			
Generates and maintains student interest	✓✓	✓	✓✓ ^{viii}
Fosters students' intrinsic motivation and responsibility for their own learning	✓	✓✓	✓
Seeks student input on course learning experience	X	✓✓	✓
Provides evidence of student learning	✓	✓✓	X

vii. "The instructor created a supportive environment that helped me learn" (ibid).

viii. "The instructor stimulated my interest in the course" (ibid).

DIMENSION OF TEACHING EFFECTIVENESS	PEER REVIEW	TEACHING DOSSIERS	STUDENT SURVEYS
DEVELOPMENT			
<i>Reflection</i>			
Reflects on and assesses teaching and learning practices	X	✓✓	X
Engages in a scholarly approach to teaching through determining and implementing best practices	X	✓✓	X
<i>Growth and Continuous Improvement</i>			
Participates in professional development activities	X	✓✓	X
Makes thoughtful and deliberate changes to practices or develops innovations in response to new information about best practices or to other opportunities as they arise	X	✓✓	X
Regularly revises or updates course content, assignments, format, or teaching strategies in response to feedback and reflection	X	✓✓	X
<i>Collaboration, Mentorship and Leadership</i>			
Demonstrates leadership related to teaching and learning within the University and in the broader educational community	X	✓✓	X
Interacts and works with colleagues around teaching and learning	X	✓✓	X
Provides and receives mentorship related to teaching, including with teaching assistants	X	✓✓	X
Contributes to the Scholarship of Teaching and Learning (SoTL) by disseminating findings from research on instructional practices	X	✓✓	X

APPENDIX

From the Course Evaluation Project Website: <https://uwaterloo.ca/waterloo-course-evaluations/frequently-asked-questions/faqs-about-pilot-test>

FAQ (general) – FAQ (pilot questions)

What Questions Were Included on the Pilot Test?

The questions included on the pilot test were developed by a subcommittee of CEPT1, after a study of the research in teaching and learning and multiple rounds of consultation with the campus community. The questions included seek to measure elements of teaching and learning recognized in the research literature and at the University to be priority areas for university instruction. The questions are also designed to avoid presuming a particular style of content delivery. Phase 2 of the project modified the instrument based on input from focus groups with students in every faculty.

The pilot questions are listed below, but we do caution that the intention is not to begin another round of consultation about the wording or selection of questions. The questions may be modified, depending on the results of the pilot testing.

The first nine questions will be measured on a 5-point Likert scale (ranging from strongly disagree to strongly agree; there will also be an additional response-category, labelled: “have no basis for rating”).

- The instructor identified the intended learning outcomes for this course.
- The intended learning outcomes were assessed through my graded work.
- The course activities prepared me for the graded work.
- Graded work was returned in a reasonable amount of time.
- The instructor helped me to understand the course concepts.
- The instructor created a supportive environment that helped me learn.
- The instructor stimulated my interest in this course.
- Overall, I learned a great deal from this instructor.
- Overall, the quality of my learning experience in this course was excellent
- The course workload demands were.... (scale ranging from very low to very high)

Additional questions, for analysis purposes of the pilot-test data, include the following:

- What is your gender identity? (note that this can also include gender expression as it relates to your gender identity).
- On average, I attend class...
- In terms of an expected grade in this course, I expect to get...
- For me, this course is (required or elective).

***For online courses, the following question: “On average, I attend class” will be replaced with: “On average, I engage in the prescribed weekly online work for this course...”