Executive Summary: Teaching Evaluation at The University of Alberta
Draft for CLE - September 2016
Prepared by the Centre for Teaching and Learning

History of Teaching Evaluation at the University of Alberta

While research is ongoing related to the history of teaching evaluation at the University of Alberta, there is evidence that student evaluation of teaching (particularly through surveys) has has been undertaken for decades. From 1978-1985, student questionnaires were individually designed and used in many departments at the University of Alberta (Kanuka et al., 2009). In 1985, an IDQ (Instructor Designed Questionnaire) system was implemented; individualized questionnaires could be created from a catalog of items acquired from the University of Michigan. In 1994, the Universal Student Ratings of Instruction (USRI) was implemented and included a common set of 14 questions. All instructors were required to utilize the USRI in all classes (Kanuka et al., 2009). Revisions to the USRI items were undertaken in 1995 and 1999; in 1999 the TLC (Teaching and Learning Committee) struck a subcommittee to investigate evaluation methods for alternate delivery courses such as web based courses, distance courses, and courses with multiple instructors. In 2001 a report from this committee was submitted to TLC by Carolin Kreber. "The committee had attempted to address the above by deriving parallel questions for various methods of delivery using a framework based on Boyer's Scholarship of Teaching" (Kanuka et al., 2009, p.34). Unfortunately, recommendations from this report were not pursued, potentially due to member turnover on TLC (Kanuka et al., 2009).

In 2009, a report titled *Evaluation of Teaching at the U of A* was put forward by a Teaching Evaluation subcommittee of the General Faculties Council Committee on the Learning Environment (CLE). This report outlined a number of recommendations related to the USRI instrument and to teaching evaluation more generally, as well as GFC policy (described in a subsequent section). Based on information located thus far, there is no evidence to indicate that the recommendations of this report were pursued.

In 2012, the CLE Chair proposed that the 2009 CLE report be revisited and that a subcommittee make recommendations based on this document. This work was undertaken by the Subcommittee on the Status of USRIs, and in 2013 the subcommittee put forward four recommendations (described in a subsequent section) and suggested that a "working group be struck to determine how to promote consistent interpretation and implementation of policy" (Andrews et al., 2013, p. 4). In addition, it was recommended that the Provost's Office move GFC policy section 111 to UAPPOL. Based on information currently available, recommendations of this report have not been pursued.

In November 2013, The Renaissance Committee, which arose through ratification by the Association of Academic Staff University of Alberta (AASUA) and the Governors of the University of Alberta, put forward a report that aimed to address aspects of the terms and conditions of work performed at the University of Alberta. The report detailed a number of concerns related to teaching and the evaluation of teaching, including USRIs, and made specific

recommendations (outlined in a subsequent section). Further, the report stated that if suggested changes to the USRIs were not completed within two years (end of Fall term, 2015), that a moratorium should be declared on their use. Based on information located thus far, there is no evidence to indicate that the recommendations of this report were pursued.

Current GFC Policy on Teaching Evaluation

The current GFC policy highlights that teaching evaluation at the University of Alberta serves two purposes: summative and formative (GFC section 111.2). Summative evaluation relates to evaluation of teaching for purposes of promotion and tenure as well as awarding excellence and withholding reward. Formative evaluation assists teachers to identify areas of strength and improvement such that they can refine their teaching skills. Further, GFC policy highlights that evaluation of teaching must be multifaceted and include the Universal Student Ratings of Instruction (USRI) as well as other methods of evaluating teaching as determined within each Faculty. This may include one or more of the following: "input from administrators, peers, self, undergraduate and graduate students, and alumni" (GFC section 111.2).

GFC policy (GFC section 111.3) outlines that the current USRI take the form of a questionnaire with ten required questions (which were last updated in 1999). Additional questions may be supplemented by faculties, departments, and instructors; students may also make comments. Since 2014, the USRI has been administered electronically by the eUSRI system (GFC section 111.3). The anonymity of student responses to the USRI "is protected under normal circumstances" but in situations where there are concerns for the safety of faculty, staff or students as a result of statements made on a USRI, the identity of the author may "be sought out and disclosed to the appropriate University officials" (GFC section 111.3) in accordance with the Protocol for Urgent Cases of Disruptive, Threatening or Violent Conduct (GFC section 91.3).

Finally, it should be noted that current GFC policy identifies a large gap in teaching evaluation at the U of A: "nowhere, in any document, is there a clear and complete statement of what constitutes excellent teaching. It is taken for granted that we all know" (GFC section 111.1). To attempt to address this gap, four principles of effective teaching and learning were outlined in section 111.1, though these were last updated in 2000. These principles relate to: the teacher as a scholar; engaging the minds of students; respect between students and teachers; and ensuring a climate conducive to learning (GFC section 111.1). While these principles provide some direction on what excellent teaching may look like, they lack specificity and fail to describe the qualities of satisfactory (or unsatisfactory) teaching (Kanuka et al., 2009). In addition, no policy or documentation exists to connect these principles to clear and measureable outcomes for evaluating teaching (Kanuka et al., 2009).

2009 Report of the Sub-Committee of the CLE: Evaluation of Teaching at the U of A (Kanuka et al., 2009)

The overall goals of this report were to examine teaching evaluation instruments, review teaching evaluation policy at the UofA and other institutions, and make recommendations for improving teaching evaluations. Each of these areas will be summarized briefly.

In part one of the report, a literature review was conducted with a focus on reviewing scholarly research related to university teaching evaluations. A total of 35 papers were reviewed, findings were organized around the following themes: validity, bias, whether students can effectively measure quality teaching, the need for effective tools, correlations between grades and ratings, the impact of instruments on quality of teaching, and evaluation of faculty for tenure and promotion (Kanuka et al., 2009). The key findings related to each of these themes are presented in Appendix A; more recent literature has also been included around each theme. Overall, the report highlighted that there is a significant body of literature related to student ratings, and that "inconsistency exists in all areas of teaching evaluation"(Kanuka et al., 2009, p.17). However, the committee stated that "among educational researchers, there is some consensus that when USRI instruments have undergone rigorous psychometric and statistical procedures it results in valid and reliable ratings"(p. 17).

The second part of the report focused on evaluation of teaching at the University of Alberta. In particular, a review of the GFC policy on teaching evaluation (GFC section 111.2) was conducted and two main points were brought forward. "First, these policy guidelines require that teaching be evaluated in a multi-faceted manner, although the particular methods of assessment are left to individual faculties. How to assess teaching is not obvious" (Kanuka et al., 2009, p. 21). It was also noted that while there were attempts to describe attributes of good teaching within the GFC document, there were no assessment methods indicated to accompany each attribute. There is also some evidence that Chairs struggled with this mandate. (See Appendix B for a description of the concerns and issues faced by Chairs.) Overall, the report highlighted the importance of providing clear direction on what constitutes excellent teaching and specific suggestions on how it could be evaluated.

The second point brought forward in the report was that "statements on teaching evaluation have not been considered by GFC for more than ten years. This means that the present policy and assessment tools, the USRI in particular, were developed before both the Dare to Discover and Dare to Deliver documents. The values articulated in these documents should be reflected in how we assess teaching at this institution" (Kanuka et al., 2009, p.22). The report emphasized that there should be coherence between values of the institution, the institutional strategic plan, the GFC policy on teaching and learning, and the evaluation process. It should also be noted that since the time of this report, a new institutional strategic plan, "For the Public Good" was approved (June 2016).

Finally, part three of the report made recommendations. The committee "concluded that a professionally developed instrument with appropriately established metrics can result in valid and reliable teaching evaluation instrument" (Kanuka et al., 2009, p.2). The committee also highlighted that "the validity and reliability of the USRI currently in use at the University of Alberta is unknown and needs to be revisited" (p.2). Specifically, the following four recommendations were made: (1) determine the purpose of the USRI. (2a) Consider the use and administration of the USRI within a broader context within course and program evaluation,

and (2b) if the USRI is to continue being used; create a professionally developed instrument leveraging expertise to ensure validity and reliability. (3) To support GFC policy, put forward a guide which includes examples for facilitating multifaceted evaluation (given that Chairs, Deans, Supervisors and Faculty struggle with this in FEC) and (4) update GFC policy (Kanuka et al, 2009, p.23). Key points from this report are summarized in Table 1.

2013 Report of the GFC Committee on the Learning Environment Subcommittee on the Status of USRIs (Andrews et al., 2013)

The purpose of the 2013 report was to summarize discussion of the subcommittee surrounding the meaning of the four recommendations in the 2009 report and how they could be actualized. With regard to recommendation 1, the purpose of the USRI, the committee determined that it is to serve both formative purposes (improve teaching) and to provide data for summative evaluation (Andrews et al., 2013, p.2). The committee advised that both the questionnaire items and student open ended comments be maintained in any future revisions to USRIs.

The committee supported recommendation 2a and stated that "USRIs should not be focused exclusively on instruction but should also include questions that shed light on the context of teaching and learning… and should have the purpose of enhancing the quality of student educational experience" (Andrews et al., 2013, p.3). The committee also suggested that when GFC policy 111 is revisited for inclusion in UAPPOL, greater clarity on procedures surrounding confidentiality and anonymity should be addressed and the "principles of good teaching/learning at the University of Alberta needs to be reaffirmed" (p.3). The committee also recommended that "a group, including internal expertise, some members of which have experience in psychometrics, should be charged to revisit questions and USRIs" (p.3).

With regard to recommendation 3 (multifaceted evaluation), the committee highlighted that "there is concern that item 'overall the instructor is excellent' is too dominant in the measure of teaching for the evaluation of instructors" (p.3). The committee recommended that a guide and / or training be made available to support multifaceted evaluation as per the requirements of GFC policy. Finally, the committee indicated that, in opposition to recommendation 4 of the 2009 report, the "GFC policy does not generally need updating" (Andrews et al., 2013, p.4). Instead, the report emphasized that consistent interpretation and implementation of the policy across faculties is required. Key points from this report are summarized in Table 1.

2013 Renaissance Committee Report (Cheeseman et al., 2013)

The Renaissance Committee arose through ratification by the Association of Academic Staff University of Alberta (AASUA) and the Governors of the University of Alberta in 2012. The report put forward by the committee aimed to acknowledge and discuss "systemic aspects of the terms and conditions of the work performed at the University of Alberta" (Cheeseman et al., 2013, p.2).

The report brought forward a number of concerns related to how teaching is carried out and evaluated at the University of Alberta. For example, while the University asserts that research

informs teaching, the Renaissance Committee found "no evidence that the University of Alberta has policies that support, measure, or reward the link" (Cheeseman et al., 2013, p.55). Further, the committee reiterated that while existing policies at the UofA highlight the importance of teaching and identify the need for multifaceted evaluation of teaching, there was "little evidence that these policies are consistently applied by faculty evaluation committees in decisions to award merit or grant tenure or promotion" (p. 44). In addition, the committee found that there is an "inappropriate use of, and reliance on, the Universal Student Rating of Instruction (USRI) as a measure of teaching effectiveness" (p.55). In particular, stakeholder discussions emphasized a frustration with the emphasis on USRIs for evaluation and found that a single item on the USRI, "the instructor was excellent" was "often the only item to find its way to FEC in an annual assessment" (p. 68). Further, the committee highlighted that "the USRI is an insufficient indicator of teaching effectiveness; that is the opinion even of those who regard the tool as reliable and valid" (p.68).

Discussion and Recommendations

   Evaluation of teaching. Based on a literature review, the report concluded that best practices in teaching evaluation "emphasize a comprehensive, clear and transparent system and extend to teaching resources, support, and recognition. An effective evaluation system has multiple sources of information, collected through a variety of methods and assessed at multiple points in time" (Cheeseman et al, 2013, p. 69). In addition, it was emphasized that evaluation of teaching should consider the effectiveness of teaching, "but also the Scholarship of teaching (rigorous examination and generation of knowledge about how specific content is best taught and how learning occurs within specific disciplines)" (p. 54). As such, teaching evaluation should examine "connections between teaching and research and how it is realized in course design, teaching approaches, and student learning" (p.54).

When considering the evaluation of teaching related to tenure and promotion (ie. summative evaluation of teaching), the committee also had specific suggestions. In particular, it advised that recommendations on teaching that are forwarded to FEC for evaluation should be undertaken by peers that are performing similar teaching duties (Cheeseman et al., 2013, p.70). Further, the committee emphasized that it is "the role of the scholar to define the value of her/his Scholarship in teaching and contextualize its standards and, if it is the case, excellence. This must be done through data and supporting narrative" (p.70).

As with teaching evaluation for formative purposes, the committee also stated that "teaching evaluations for the purpose of tenure and promotion must be multifaceted" (p.70). Sources of data on teaching could include:

   student ratings of courses, a teaching dossier, peer observations, external reviews of content, reflection of the teacher (self-assessment), administrator reviews of content and course observation, review of published work on teaching Scholarship, and evidence supporting the reputation of the teacher in the field(s) of instruction, within and without the University. (p.70)

Further, the committee highlighted that evaluation can go beyond the act of teaching and also incorporate course and instructional design as well as implementation, and results. Thus, evaluation of teaching should reflect the "full complement of teaching duties of the faculty member" (p.70).

Feedback from Students and USRIs. It was the position of the committee that "feedback from students is an essential component of teaching evaluation" (p. 68). While student feedback can be accomplished through a variety of means including samples of student work, interviews, and standardized (national) exams; "formal course evaluations are the typical format for student feedback on teaching effectiveness"(p. 68). Related to USRIs, the report emphasized that there is:

> *ample evidence that, when used appropriately and as a part of a broad evaluation of teaching, student evaluations can provide a valid measure of teaching effectiveness. However, the literature also highlights many ways that bias arises in the use of student evaluation questionnaires that have flawed design and implementation.* (p. 68-69)

Through meetings with stakeholders, the committee suggested that the University of Alberta design a set of questions on the USRI that evaluate the effectiveness of teaching, and that this work should be undertaken in partnership with the Centre for Teaching and Learning and student representatives, and completed within two years. In addition, the report highlighted that the question of anonymity on USRIs be addressed: "it is imperative that students take ownership of their responses. Doing so would lead to fewer – perhaps no – cases of abusive, unethical, sexist, or threatening language" (p. 68)*.* It was emphasized that the University must protect the wellbeing of its teaching staff as well as the identity of students, and could do so "by creating a set of questions that would no longer require a freeform section on the evaluation, removing the occasion for abusive statements" (p.68). Specific recommendations of the report related to USRIs and to Teaching Evaluation are included in Appendix C. Key points from this report are also presented in Table 1.

Table 1: Summary of Positions and Recommendations Related to the Evaluation of Teaching in University of Alberta Policy, Documents and Reports

Teaching Evaluation	
Position / Recommendation	**Supported by**
Teaching evaluation should be multifaceted.	GFC Policy 111.2 2009 Report of Sub-committee of CLE (Teaching Evaluation). 2012 Association of Academic Staff University of Alberta (AASUA) Position Statement on USRIs. 2013 Report of Sub-committee of CLE (Status of

	USRIs)
Chairs, Deans, Supervisors and Faculty may struggle with implementing and undertaking multi-faceted evaluation, and need support in doing so.	2009 Report of Sub-committee of CLE (Teaching Evaluation).

2013 Report of Sub-committee of CLE (Status of USRIs) |
| A multifaceted teaching evaluation plan / guide / strategy should be developed. This should include definitions, examples, and specific strategies for implementation and training. | 2009 Report of Sub-committee of CLE (Teaching Evaluation).

2013 Report of Sub-committee of CLE (Status of USRIs)

2013 Renaissance Committee Report |
| FEC decisions regarding promotion and tenure must be based on multiple indicators of teaching (i.e. be multifaceted). This may not have been consistently applied in the past. | GFC Policy 111.2

2009 Report of Sub-committee of CLE (Teaching Evaluation).

2012 Association of Academic Staff University of Alberta (AASUA) Position Statement on USRIs.

2013 Renaissance Committee Report - *this could include "student ratings of courses, a teaching dossier, peer observations, external reviews of content, reflection of the teacher (self-assessment), administrator reviews of content and course observation, review of published work on teaching Scholarship, and evidence supporting the reputation of the teacher in the field(s) of instruction, within and without the University*" (p.70). |
| Evaluation of teaching should not focus solely on course-based evaluation; broader teaching duties should also be included. (For example, graduate student supervision and mentoring, course design, curriculum development, etc). | 2013 Renaissance Committee Report |
| Peer review should be a part of evaluation for tenure and promotion. | 2013 Renaissance Committee Report |

Opportunities for teacher training and support are needed.	2012 Association of Academic Staff University of Alberta (AASUA) Position Statement on USRIs. 2013 Renaissance Committee Report
Research should inform teaching; currently U of A policies do not support, measure, or support this link.	2013 Renaissance Committee Report
Student input should be sought in teaching evaluation, potentially through USRIs or similar instruments.	2009 Report of Sub-committee of CLE (Teaching Evaluation) 2013 Renaissance Committee Report
There are aspects of teaching that students cannot evaluate.	2009 Report of Sub-committee of CLE (Teaching Evaluation) 2012 Association of Academic Staff University of Alberta (AASUA) Position Statement on USRIs.

USRIs	
Position / Recommendation	**Supported by**
Purpose of USRI must be determined.	2009 Report of Sub-committee of CLE (Teaching Evaluation) 2013 Report of Sub-committee of CLE (Status of USRIs) - *current purpose is both formative and summative.*
Open-ended comments on USRIs: ● should be included	 2013 Report of Sub-committee of CLE (Status of USRIs) - *comments provide valuable formative feedback.*
● should not be included	2013 Renaissance Committee Report - *comments introduce the occasion for abusive statements*
● student identities should not be included in reports to instructors but kept on record (for the protection of instructors and students).	2012 Association of Academic Staff University of Alberta (AASUA) Position Statement on USRIs. 2013 Renaissance Committee Report

Use and administration of USRI must be considered in broader context (not just focused on teaching).	2009 Report of Sub-committee of CLE (Teaching Evaluation) 2013 Report of Sub-committee of CLE (Status of USRIs). - *focus should not be exclusively instruction; questions should be included to shed light on the teaching and learning context.*
The current USRI instrument is outdated, lacks validation, and needs redevelopment.	2009 Report of Sub-committee of CLE (Teaching Evaluation) 2013 Report of Sub-committee of CLE (Status of USRIs) 2013 Renaissance Committee Report
The *required* USRI items need to be modified to apply to multiple teaching contexts. Additional (optional) question variants should be developed that apply to diverse teaching contexts, including but not limited to: lectures, lab settings, online or blended courses, other IT enhanced courses, clinical settings, problem based/context based teaching.	2002 Teaching and Learning Committee (TLC) Report from the Sub-committee on Evaluation of Alternate - Delivery Courses
A professionally developed instrument should be created by an expert / experts to ensure validity and reliability.	2009 Report of Sub-committee of CLE (Teaching Evaluation) 2013 Report of Sub-committee of CLE (Status of USRIs) - *"a group, including internal expertise, some members of which have expertise in psychometrics, should be charged to revisit questions and USRIs" (p.3).* 2013 Renaissance Committee Report
A moratorium on USRI use should be implemented until redevelopment occurs (deadline: end of Fall Term, 2015).	2013 Renaissance Committee Report
USRIs should be used as part of a broader teaching evaluation; they should **not** be the sole measure of teaching performance. [While this is outlined in GFC policy, it may not be utilized in practice.]	GFC Policy 111.2 2009 Report of Sub-committee of CLE (Teaching Evaluation) 2012 Association of Academic Staff University of

	Alberta (AASUA) Position Statement on USRIs. 2013 Renaissance Committee Report
Concern over specific USRI items: ● Concern that "the instructor was excellent" was the only item used in FEC assessments.	2013 Report of Sub-committee of CLE (Status of USRIs) 2013 Renaissance Committee Report

GFC Policy	
Position / Recommendation	**Supported by**
GFC Policy should be revisited for inclusion in UAPPOL	2013 Report of Sub-committee of CLE (Status of USRIs)
GFC Policy revisions: ● Revisions required	2009 Report of Sub-committee of CLE (Teaching Evaluation). *"Quite simply, existing policy is in need of updating" (p.2).* 2013 Renaissance Committee Report - *called for review and update of teaching and learning policies currently in place in the GFC policy manual*
● Revisions not required	2013 Report of Sub-committee of CLE (Status of USRIs). - *policy does not generally need updating; "there is ample clarity in the existing policy as to what USRIs entail and should be used for. What is needed is consistent interpretation and effective implementation across all faculties in a way that accomplishes the multiple purposes of USRIs" (p.4).*

Multifaceted Evaluation of Teaching: Recommendations

Existing policy at the University of Alberta requires that evaluation of teaching be multifaceted (GFC section 111.2), though how this has been implemented at the university has not always been clear (Kanuka et al., 2009). In 2015, the Centre for Teaching and Learning (CTL) hosted a symposium on Multifaceted Summative Evaluation of Teaching, wherein a number of recommendations for best practice were brought forward. The key points from this symposia, along with themes present in scholarly literature, U of A documents and reports (see Table 1 and Table 2), and the practices of other institutions will be summarized briefly here.

Defining Effective Teaching

As part of multifaceted evaluation of teaching plan, the University of Alberta must define effective teaching. There must be a specific set of criteria that can be used for purposes of evaluation; these criteria should be shared with faculty, instructors and students.

Formative and Summative Evaluation

Both formative and summative evaluation of teaching should be multifaceted. Multiple sources of evidence collected at multiple times should be used to evaluate teaching quality for purposes of tenure and promotion (summative evaluation) to ensure fair, thorough, and valid evaluation. Multiple sources of feedback also provide detailed insight into specific areas of strength or improvement related to different aspects of teaching (i.e. planning and organization, graduate student supervision, communication, grading, etc).

Implementation Plan Including Support and Education

A multifaceted teaching evaluation plan /strategy should be developed. This would detail how existing U of A policy can be implemented and should include definitions, examples, evaluation procedures, and specific strategies for training and support. It should reflect U of A values and policy, be consistent with the evaluation plans of other institutions, and be grounded in scholarly literature.

Table 2: What should a multifaceted teaching evaluation plan include?

Feature	Support
Detailed description of effective (excellent, good) teaching. This should include classroom instruction as well as the scholarship of teaching, preparation, planning, organization, and other aspects of teaching.	2013 Report of Sub-committee of CLE (Status of USRIs). 2013 Renaissance Committee Report
Explicit connection between the qualities of effective teaching and the measures used to evaluate teaching. (e.g. items on USRI and other measures should align with qualities of effective teaching).	2009 Report of Sub-committee of CLE (Teaching Evaluation).
Documents that support consistent implementation, use, and interpretation of USRIs (if they are used). This includes provision for formative feedback.	2013 Report of Sub-committee of CLE (Status of USRIs). 2013 Renaissance Committee Report
Documents that support implementation, use, and interpretation of other methods	2013 Renaissance Committee Report

of teaching evaluation for both summative and formative purposes. (i.e. Peer review, portfolios, etc).	
Provision for teaching development and training. (Support, scholarship). This should be partnered with teaching evaluation.	2013 Renaissance Committee Report

Evidence of teaching within a multifaceted evaluation plan may involve:

Student Evaluations of Course Instruction (USRIs)

Based on the review of policy and documents at the University of Alberta, there is a general consensus that student input should be sought related to their experience with course instruction and the learning environment, potentially through USRIs or similar instruments (see Table 1). This is also evident in the practices of other institutions (see Appendix D). For example, Stanford University recently (Autumn Quarter 2015-16) introduced a new end of term course evaluation which aimed to "focus on learning, increase student self-reflection, provide instructors and departments with useful data, be customized and relevant to each course, and support broad Stanford priorities: using feedback to improve student learning, courses, and programs, and using analytics to address critical questions" (Stanford University VPTL, *Stanford's New Course Evaluations,* n.d.). The instrument includes nine required items and additional customizable questions, which may be open-ended or closed-ended (Stanford University VPTL, *Standard Course Evaluation Questions*, n.d.). For more information on the Stanford Course Evaluations, or to view specific items, go to Appendix D.

Rather than adopt a single instrument to obtain student feedback, some institutions are utilizing multiple instruments to seek insight on students' perceptions of teaching and learning as well as the broader context of the student experience. For example, both the University of Oxford and the University of Sydney have recently adopted the Student Barometer as a broad measure of student experience, including learning experience, living experience, support services, and other areas. This measure is administered once per year and aims to "track and compare the decision-making, expectations, perceptions and intentions of students from application to graduation" (University of Sydney, 2016a, para. 2). Specific items from the Student Barometer can be found in Appendix D. In addition to this instrument, the University of Oxford also employs local mechanisms for evaluation within departments (University of Oxford, 2015, p. 7) as well as the National Student Survey for undergraduate students in the last year of their program.

The University of Sydney also employs a Student Experience Survey (University of Sydney, 2016b) for undergraduate students in their first and final year of their program and a mandatory Unit of Study Survey (course survey), which is similar to a USRI. The Unit of Study Survey (USS), which has been recently updated, is administered online and contains 8 required items

(6 quantitative, 2 open response) and up to 4 faculty-specific quantitative items and one faculty-specific open response item. Each faculty can also have up to 4 USS versions to allow customisation of the survey for different contexts (University of Sydney, 2016c). (see Appendix D or the University of Sydney website for an example of the USS survey).

Taken together, the examples provided here highlight that other institutions value student feedback on the teaching and learning environment and are making efforts to update and improve the instruments they utilize to obtain this feedback. While the current USRI employed at the University of Alberta requires updating (Cheeseman et al., 2013; Kanuka et al., 2009), it is not the recommendation of this summary that the U of A adopt an instrument from another institution outright. Further investigation is required related to the effectiveness of these instruments and how they have been developed, validated, and employed. In addition, prior to adoption, there must be consideration of how data from instruments such as these can be used to inform and improve teaching and learning environments at the U of A within the context of a multifaceted teaching evaluation plan.

Peer Review Of Teaching / Classroom Observations

Peer review of teaching is the "informed collegial assessment of faculty teaching for either fostering improvement or making personnel decisions" (Gibson, n.d., para 5) and can contribute to multifaceted evaluation of teaching. In his keynote speech at the 2015 CTL Multifaceted Summative Evaluation of Teaching Symposia, Dr. Richard Price discussed the implementation of a Peer Review Teaching (PRT) Initiative at the University of British Columbia. This PRT process was implemented as part of multifaceted summative evaluation of teaching for the purpose of judgment for promotion and tenure. The key recommendations highlighted by Dr. Price related to the implementation of summative PRT were:

1) *To balance teaching context and consistency in PRT.* This meant acknowledging that peer review may need to be adapted to specific contexts, but that there should also be consistency in the products of peer review to ensure its fair to all staff.
2) *To develop training plans to prepare staff to engage in PRT.* This included training (workshops) to assist faculty to engage in summative PRT, but to also foster an environment that supported formative peer review as well.
3) *To outline the elements of summative PRT.* The five elements determined by the committee at UBC (and widely adopted by faculty) were:
   a) Evidence - this involves defining what data is collected. This should include peer observation but may also include student evaluations, syllabi, statements of teaching practice, evidence of supervision, and other data sources.
   b) Standards - this involves paralleling the PRT process with research assessment processes, which may be accomplished through the use of one 'external reviewer' (outside the department or unit) and one 'internal reviewer'. In addition, reviewers were not asked to make recommendations on promotion or tenure; rather, they made assessments regarding whether standards of teaching were exceeded, met, or not met. Feedback was qualitative and narrative in nature;

numerical scores were not utilized within the Faculty of Arts at UBC nor were they recommended by Dr. Price.

c) Criteria - this involves defining the qualities of good teaching that will be assessed during PRT. A specific list or framework is recommended. Some potential examples are [Sample Observation Tool (Checklist)](#) and the [Sample Observation Tool (Likert)](#).

d) Ethics - this involves issues related to confidentiality, professionalism, and legality. The specific legal status of reports, emails, and other documentation related to the PRT process should be clearly outlined. It was the recommendation of Dr. Price that a team-written report be created by the two reviewers and that this report be made available to the candidate.

e) Protocol - involves the scheduling and implementation of peer review. For promotion and tenure purposes, Dr. Price recommended one formative peer review, one summative peer review at approximately 3 years, and one summative review just prior to recommendations for promotion and tenure.

Additional information and resources related to Peer Review of Teaching (summative and formative) can be found on the [CTL website](#).

Teaching Dossiers / Portfolios

A teaching dossier is a "document intended to facilitate the presentation of a faculty member's teaching achievements and major strengths for self-assessment and interpretation by others" ([Day, Robberecht & Roed, 1996, p.1](#)). Teaching dossiers are a cumulative record of one's teaching activities and often include: "(a) a statement regarding the faculty member's teaching philosophy, goals, and strategies; (b) a description of teaching (planning, preparing, and teaching courses; assessing student learning; and giving feedback); (c) an evaluation of teaching accomplishments; and (d) suggestions regarding possible changes for future teaching" (Day et al.,1996, p.1). Teaching dossiers or portfolios can play an important role in the multifaceted evaluation of teaching because they require instructors to gather multiple sources of evidence and define the value of their scholarship in teaching (Cheeseman et al., 2013). Related to summative evaluation of teaching, the 2013 Renaissance Committee Report recommended that "a teaching dossier, following CTL standards,should be part of all tenure and promotion packages" (Cheeseman et al., 2013, p.70). More information and resources related to teaching dossiers are available on the [CTL website](#).

Other Evidence of Teaching

A number of other sources of evidence could be used to evaluate teaching. These may include but are not limited to "*external reviews of content, reflection of the teacher (self-assessment), administrator reviews of content and course observation, review of published work on teaching Scholarship, and evidence supporting the reputation of the teacher in the field(s) of instruction, within and without the University*" (Renaissance Committee Report, 2013, p.70). Other potential sources of evidence for effective teaching (which may be included in a teaching portfolio or evaluated separately) have been suggested by the [University of Sydney](#).

References

Andrews, N., Chelen, D., Connor, B., Kostiuk, L., Kwong See, S., Milner, R. (2013). *Report of the GFC Committee on the Learning Environment Subcommittee on the Status of USRIs.* Retrieved from: https://www.ualberta.ca/-/media/ualberta/centre-for-teaching-and-learning/university-governance/item-5-usri-subcommittee-final-report-june-2013-final.pdf

Association of Academic Staff University of Alberta (AASUA). (2012). *AASUA Position Statement on URSIs.* Retrieved from: http://www.aasua.ca/wp-content/uploads/2014/03/AASUA_Position_Statement_on_USRIs.pdf

Centre for Teaching and Learning (CTL). (2015). *Multifaceted Summative Evaluation of Teaching Symposium.* Retrieved from: https://www.ualberta.ca/centre-for-teaching-and-learning/teaching-and-learning-events/symposium-series/past-symposia/multifaceted-summative-evaluation-teaching

Cheeseman, C., MacLaren, I., Carey, J., Glanfield, F., Liu, L., McFarlane, L., Cahill, JC., Garneau, T., Supernant, K., & Szeman, I. (2013). *Report of the Renaissance Committee.* Retrieved from: http://www.renaissance.ualberta.ca/~/media/renaissance/RenComReportComplete20131209Footnote.pdf

Day, R., Robberecht, P., & Roed, B.(1996). Teaching Dossier: A guide. Retrieved from: https://d1pbog36rugm0t.cloudfront.net/-/media/ualberta/centre-for-teaching-and-learning/instructional-resources/teaching-dossier/teachingdossierguide-1.pdf

Erkut, E. & Kreber, C. (2002). *Report from the Sub-committee on Evaluation of Alternate-Delivery Courses: Continuing Discussion.* General Faculties Council Teaching and Learning Committee.

General Faculties Council. *GFC Policy Manual Section 111: Teaching and Learning and Teaching Evaluation.* Retrieved from: http://www.gfcpolicymanual.ualberta.ca/en/111TeachingandLearningandTeach.aspx

General Faculties Council. *GFC Policy Manual Section 91: Protocol for Urgent Cases of Disruptive, Threatening or Violent Conduct.* Retrieved from: http://www.gfcpolicymanual.ualberta.ca/91ProtocolforUrgentCasesofDisr.aspx

Gibson, S. (n.d.) *University of Alberta Peer Review of Teaching.* Retrieved from: https://www.ualberta.ca/centre-for-teaching-and-learning/teaching-and-learning-events/s

ymposium-series/past-symposia/multifaceted-summative-evaluation-teaching/peer-revie
w-of-teaching

Kanuka, H., Marentette, P., Braga, J., Campbell, K., Harvey, S., Holte, R., Nychka, J., Precht,
D., Read, D., Skappak, C., & Varnhagen, C., (2009). *Evaluation of Teaching at the U of
A: Report of the Sub-Committee of the Committee on the Learning Environment (CLE).*
Retrieved from:
https://www.ualberta.ca/-/media/ualberta/centre-for-teaching-and-learning/university-gov
ernance/cle-evaluaton-of-teaching-at-the-university-of-albertajanuary2009.pdf

Price, R. (2015). A*chievements and Challenges of UBC's Peer Review of Teaching Initiative: A
View from the Faculty of Arts* [Video file]. Keynote speech at the Centre for Teaching and
Learning Multifaceted Summative Evaluation of Teaching Symposium. Retrieved from:
http://livestream.com/ualberta/msets-05-09-15

Stanford University Vice Provost for Teaching and Learning (VPTL). (n.d.). *Standard Course
Evaluation Questions*. Retrieved from:
https://vptl.stanford.edu/teaching-learning/teaching-practices/evaluation-feedback/stanfo
rds-new-course-evaluations/standard

Stanford University Vice Provost for Teaching and Learning (VPTL). (n.d.). *Stanford's New
Course Evaluations.* Retrieved from:
https://vptl.stanford.edu/stanfords-new-course-evaluations

University of Oxford. (2015). *Procedures for the Annual Monitoring of Courses.* Retrieved from:
http://www.admin.ox.ac.uk/media/global/wwwadminoxacuk/localsites/educationcommitte
e/documents/qualityassurance/Procedures_for_annual_monitoring_of_courses_FINAL_
Oct15.pdf

University of Sydney. (2016a). *Student Barometer (SB/IB)*. Retrieved from:
http://sydney.edu.au/education-portfolio/ei/studentbarometer/

University of Sydney. (2016b). *Student Experience Survey (SES)*. Retrieved from:
http://sydney.edu.au/education-portfolio/ei/ses/

University of Sydney. (2016c). *Unit of Study Survey (USS)*. Retrieved from:
http://sydney.edu.au/education-portfolio/ei/USS/default.htm