

REPORT OF THE COURSE EVALUATION PROJECT TEAM

November 7, 2016

Introduction

In 2014, the Associate Vice-President, Academic established the Course Evaluation Project Team (CEPT) to explore the potential for a new, campus-wide, course evaluation model. For the past two years, the project team has reviewed the literature on course evaluation and conducted consultations across campus with representative stakeholders regarding the possible implementation of a new course evaluation model. Over the next several months, the consultation and decision-making process will provide further opportunities for review and suggestions by the campus community. The project team will then submit a revised report to the Vice President, Academic and Provost for a decision about next steps.

The review of literature and consultation process has culminated in a number of proposed recommendations, which are explained in this report. At this critical stage, the project team seeks feedback from all campus stakeholders in response to the following recommendations:

1. Adopt a cascaded course evaluation model that includes a common set of standard questions across all Faculties, with optional additional questions chosen by each Faculty, academic unit, and instructor from an established question bank.
2. Develop course evaluation questions that align with the design principles outlined in this report (see section 2.0).
3. Provide access to the numeric data results from the common questions to all members of our university community.

Specific questions on these recommendations are included below in Section 7.0 of this report.

The project team recognizes the limitations of course evaluations while also acknowledging that course evaluations serve an important function for university operation and success. Data generated by course evaluations represent one of several sources of evidence to be considered for promotion and tenure, and for annual performance review purposes. Since student feedback is only one data source, other sources of evidence of teaching/course quality should take on a substantially enhanced role (see [Policy 77](#)). There is a need for continued discussion about how methods such as peer evaluation, teaching dossiers and other approaches could be applied in a consistent, systematic manner campus-wide to evaluate both teaching and course design and delivery. In order for course evaluations to be credible sources of information, any instruments used should be validated and the data recognized as student perceptions of teaching effectiveness and the learning experience in a course.

This report describes the project mandate and process, the cascaded course evaluation model, the guiding principles for instrument design, sample questions, course evaluation processes, and the technology, implementation and management of the course evaluation system. The report explains the context for recommendations and associated processes. The report concludes with next steps and specific questions for stakeholder feedback.

1.0 Project Mandate and Process

The Course Evaluation Project Team was established to provide recommendations regarding the design and implementation of a student course evaluation model. This initiative arises from a commitment to enhance the quality of the educational experience for Waterloo's undergraduate and graduate students by providing instructors with feedback about the effectiveness of their course design and delivery.

Specifically, the project team was mandated to accomplish the following:

1. Examine the various administrative, logistical, technological, and cultural issues pertaining to course evaluations at the University of Waterloo
2. Establish best practices concerning all aspects of course evaluations based on a review of the literature
3. Consider the implications of adopting changes to current course evaluation procedures in relation to Policy 77, the [MOA](#) (Memorandum of Agreement with the Faculty Association), and faculty annual performance evaluations
4. Assess the feasibility of designing a common institutional survey instrument, with customizable sections at the Faculty, department, or instructor level (referred to here as a "cascaded" model)

The project team is composed of representatives from the major stakeholder groups at the University: faculty representation (academic Faculties; Faculty Association of the University of Waterloo – FAUW); undergraduate students (Federations of Students - FEDS); graduate students (Graduate Students Association – GSA); academic support units (Centre for Extended Learning – CEL; the Centre for Teaching Excellence – CTE); and the University's Information Systems and Technology group – IST.

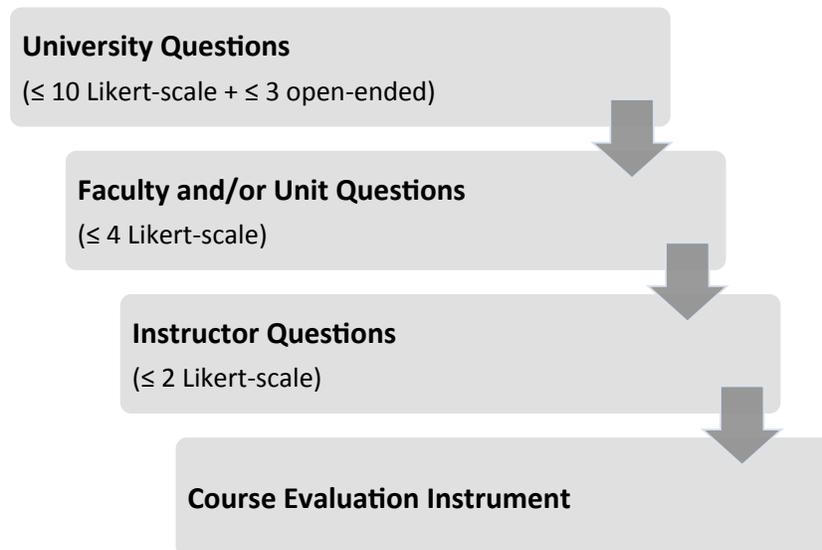
The project team and its subgroups have met regularly since May 2014. Recommendations have been informed by the review of appropriate literature, consultations with colleagues at other universities, and the review of a number of peer university websites to identify best practices and factors to consider when designing, implementing, and interpreting course evaluations. In addition, the team has carefully considered perspectives and advice offered by the university's AccessAbility Services (AAS), the Office of the President (Special Advisor on Women's and Gender Issues), as well as subject matter specialists, including social psychologists, survey design methodologists, and teaching fellows at Waterloo. A consultation process was undertaken throughout 2015 with the Senate, Deans Council, FAUW, FEDS, GSA and all six Faculties (Applied Health Sciences, Arts, Engineering, Environment, Mathematics, Science).

Three key concerns emerged from these consultations: (1) inherent biases in evaluation, (2) the advisability of university-wide questions, and (3) the privacy of and access to data. The recommendations in this report address these concerns.

2.0 The Utility of a Cascaded Course Evaluation Model

The project team recommends the adoption of a cascaded course evaluation model. This multi-layer model includes a common set of standard questions across all Faculties, complemented by optional additional questions chosen by each Faculty, academic unit, and instructor from an established question bank. (See Figure 1).

Figure 1. Cascaded Course Evaluation Model



Current course evaluation practices and instruments are varied at Waterloo. The adoption of a common set of university-wide course evaluation questions would enable us to report institutionally on a key component of our mission – student perceptions of teaching.

Institutional reporting is fully consistent with the growing expectation from government and the public for transparency and accountability from our post-secondary, publicly-funded universities. The Ontario Undergraduate Student Alliance has called for increased student access to course evaluation data, and Waterloo's Federation of Students has publicly advocated for access to aggregate data. A December 2015 report by the Ministry of Training, Colleges, and Universities (MTCU – renamed “Ministry of Advanced Education and Skills Development” in 2016 - identified course evaluation data in its list of additional vectors that could be used in advancing an outcomes-based funding model for post-secondary education in Ontario.ⁱ

Having university-wide questions is common at other universities, as noted in a recent survey done for a MTCU-funded research project on evaluations of teaching. The researchers found that 94% of the Ontario universities surveyed (n=20) had institution-wide student evaluations of teachingⁱⁱ. In addition, Canadian universities of comparable size and prominence have already moved to a cascaded course evaluation model (e.g., Toronto, McGill, UBC, Simon Fraser).

This cascaded model enables universities to have data for reporting and gives Faculties, departments, and individual instructors the ability to select additional questions for more customized feedback. Moreover, different additional questions may be selected over time, such as when instructional practices are changed in a course.

3.0 Course Evaluation Design

The project team proposes a course evaluation model structured on a set of guiding principles. The primary principle is that the course evaluation questions need to connect to a well-grounded, empirically informed definition of effective teaching.

The project team's review of research into the elements of effective teachingⁱⁱⁱ shows that effective instructors design and deliver courses that result in meaningful student learning^{iv}. While course evaluations do not measure student learning (that is the role of tools such as assignments, tests, and exams), students can provide useful feedback about how well the design and delivery of a course facilitated their learning (or not) and affected their learning experience. This understanding - together with course evaluation instruments^v used elsewhere and current literature about such instruments^{vi} - allowed the project team to identify three main dimensions to be evaluated in the instrument: Course Design, Course Delivery, and Learning Experience.

An analysis of Waterloo's existing course evaluation instruments has revealed that few questions explicitly focus on the student learning experience. The questions also privilege lecture-based instructional practices^{vii}, and there is considerable variety in the wording and number of questions asked. Instructional and assessment practices have shifted over the past few decades to embrace an expanded repertoire of options (e.g., collaborative learning, active- and problem-based learning, authentic assessments). There has also been increased focus on learning outcomes and the use of educational technologies^{viii}. Course evaluations that capture this evolution in course design, delivery, and the student learning experience are considered necessary and valuable.

The project team recommends using the following principles to guide the development of course evaluation questions:

1. course evaluations should focus on students' perceptions of the quality of course design, course delivery, and the learning experience
2. course evaluations should provide instructors with helpful, timely student feedback
3. unit chairs/directors should be able to use trends evident in successive course evaluations as one means to help ensure high quality teaching for their academic programs
4. results from scaled questions should be viewed as student perceptions of effective teaching and their learning experiences that will be further illuminated by open-ended comments
5. the selection of indicators of effective teaching and the wording of instrument items should be guided by the research literature as well as by ongoing assessment of evaluation instruments
6. evaluation questions should focus on instructional elements that students can reliably evaluate and avoid ones they cannot reliably evaluate^{ix}
7. institution-wide questions should transcend course delivery modalities and disciplines, and
8. the instrument should allow for the assessment of diverse teaching approaches with a combination of open-ended and Likert scale questions.

Guided by these principles, the project team drafted sample questions for each of the three dimensions of effective teaching (see Table 1). These are examples of the types of questions that could be used for the core, institution-wide portion of the cascaded course evaluation model. Most of the questions would be answered on a Likert scale; open-ended questions are marked with an asterisk.

As context, the *Course Design Dimension* seeks feedback regarding how well the intended learning outcomes, assessments, and course activities are aligned (good alignment should facilitate learning); the *Course Delivery Dimension* seeks feedback on how well the elements of delivering a course support student learning; and the *Learning Experience Dimension* seeks feedback on the student learning experience in the course.¹

Table 1. Dimensions of effective teaching and sample questions

Course Design Dimension	Course Delivery Dimension	Learning Experience Dimension
<ul style="list-style-type: none"> • I knew what I was expected to learn in this course • The graded work assessed what I was expected to learn • The course activities prepared me for the graded work • The coursework demands were...(Likert scale answer choices will reflect workload intensity – for example, very light to very heavy) 	<ul style="list-style-type: none"> • The instructor returned graded work in a reasonable amount of time • The instructor was a clear communicator • The instructor created a supportive environment that helped me learn • The instructor stimulated my interest in this course 	<ul style="list-style-type: none"> • The most important thing I learned in this course was* • Overall, I learned a great deal from this instructor • Overall, the quality of my learning experience in this course was excellent • What helped me to learn in this course?*
<ul style="list-style-type: none"> • What changes, if any, would I suggest for this course?* 		

*Denotes an open-ended question

4.0 New and existing model use

Waterloo’s [Policy 77](#) states that “student evaluations are an important source of information” in the assessment of teaching. However, teaching at Waterloo is assigned to a broader community than faculty and therefore course evaluation processes need to consider the entire instructional community. “Instructor”, as used in this report, includes all tenured and tenure-track faculty, adjunct appointments, lecturers, sessional instructors, and teaching assistants (TAs) who are in independent instructional roles.

Tenured professors and continuing instructors would use the new evaluation model. Instructors whose start date is after the commencement of a new evaluation process should be assessed with the new evaluation instrument. Faculties should offer instructors whose start date was prior to the commencement of a new course evaluation process the option to be assessed with (a) the previous Faculty-based instrument or (b) the new campus-wide instrument, until they have been awarded tenure and promotion to associate professor, or attained continuing status.

¹ More information about the research basis for the three dimensions and sample questions is available in a document posted on the project website. A comparison of the sample questions to Waterloo’s existing course evaluation questions can also be found on the website.

5.0 Supportive Software Platform

The project team recognizes the benefits of online delivery of course evaluations. These include:

1. lowering resource costs when compared with paper-based approaches
2. easing the work to analyze, share and post data
3. adding flexibility to accommodate a cascaded evaluation model
4. increasing security of student access; and
5. enhancing accessibility by campus stakeholders to the evaluation process and its outcomes.

A locally developed online delivery system – *eEvaluate* – has been under development in the Faculty of Science for several years. Five of our six Faculties (Applied Health Sciences, Engineering, Environment, Mathematics and Science) have fully deployed *eEvaluate* using their existing instruments, and the Faculty of Arts has conducted an extensive pilot. Faculties have reported that *eEvaluate* has largely met expectations, and this software has effectively become the *de facto* campus solution. The project team therefore concludes that the most reasonable and beneficial course of action would be for all Faculties to adopt *eEvaluate*.

An advisory committee has been struck to provide input to priorities for the technical development of *eEvaluate*. This committee is accountable to the office of the Associate Vice-President, Academic (AVPA) and the Associate Provost, Graduate Studies (APGS), and provides reports to the University Committee on Information Systems and Technology (UCIST). If the cascaded model of the instrument is adopted, a full review of software requirements to support the model will be initiated. The technical advisory committee would assist with that review.

6.0 Management of the Course Evaluation System

This section of the report addresses several high-profile issues that have been raised in project team discussions and through discussions with campus stakeholders.

6.1 The Issue of Bias

The team recognizes that every opportunity must be taken to enhance clarity of question intent, and to minimize the potential for inappropriate comments. Research acknowledges that socio-cultural variables, biases, the "halo effect" and other influences can affect course evaluation results^x. For example, student participation in course evaluation exercises can be compromised by factors such as bias (e.g. gender and race) in perceptions of course and instructional quality; indifference to the exercise by students and/or instructors; immaturity of respondents; misunderstandings of the purpose and application of course evaluation results; and instrument questions that are inappropriate or simply cannot be answered in an informed manner by students, among other factors and variables.

Institutional and individual bias regarding specific groups is a challenge that we face in our society and in higher education. There is no question that biases (e.g. sexism, racism, ageism) exist on any campus, and that these biases can be expressed in course evaluations. For example, in a study done at the University of Waterloo, when students received low grades, they gave statistically lower overall ratings of quality (course and instructor quality ratings were combined) to female instructors than to male instructors.^x Similarly, we note that bias can be a factor in the interpretation of course evaluation data by university administrators (e.g. academic chairs/directors, staff). These are serious issues that the project team has discussed and addressed.

6.2 Capacity Building of Course Evaluation Users

It is not possible to control individual behaviours and responses to course evaluation questions. However, it is possible to reduce the potential for bias, in its many potential forms, through the design of the course evaluation instrument. While it is impossible to anticipate every potential factor that could compromise the quality, validity and fairness of evaluation responses, a properly designed and implemented training and orientation program can enhance the utility of course evaluations.

Many universities have designed and implemented training and education programs for students, staff and instructors to support and guide the course evaluation process. Accordingly, we recommend that all University students, faculty, staff (Faculty and departmental) administrators and system administrators are trained in, and oriented to, the course evaluation instrument and the use and interpretation of results in addition to orientation to the *eEvaluate* platform. Training and orientation content should comprise a generic core of information, plus material that meets the information needs of specific evaluation users. These information needs should be determined following consultation with each course evaluation user group.

With regard to the course evaluation instrument, training and orientation content must address issues such as the intent of course evaluation, how and why course evaluations are used, how to interpret the results, the need to acknowledge the importance/role of bias (especially concerning gender and race) when completing and interpreting evaluations, and ethical obligations generally. In terms of the *eEvaluate* platform/technology, training and orientation content should explain the key features of the *eEvaluate* system and provide links to useful on-line resources (e.g. FAQs, instructional videos) that meet the needs of different user groups.

Training and orientation content should be accessible by course evaluation users on an “on demand” basis via a single, dedicated on-line course evaluation portal which would also enable access to the *eEvaluate* course evaluation instrument and useful resources. Mandatory training and orientation content should be presented as a “toolkit,” with online sub-sites dedicated to specific course evaluation user group information needs.

6.3 Testing, Monitoring and Evaluation: Instrument and Toolkit

In considering the adoption of a cascaded course model and implications for Policy 77, the project team recognizes that quality measures of the course evaluation instrument are needed. The testing of course evaluation instrument results will determine the reliability and validity of the instrument, including the influence of variables that could bias results at Waterloo. The results of this testing should be used to revise the instrument and/or the educational toolkit as appropriate before and following implementation to determine the influence of such variables at Waterloo.

Refinements to the course evaluation instrument should be made as necessary, following consultation with key campus stakeholders (including FAUW, GSA, and FEDS) and regular expert review of operations and instrument design and performance. Further, a full assessment of the course evaluation instrument and platform should take place after 3-5 years of campus-wide application, with findings reported to Senate.

The Office of the Associate Vice-President, Academic (for undergraduate courses) and the Associate Provost, Graduate Studies (for graduate courses) should be responsible for oversight, coordination and reporting of campus-wide course evaluation through the Quality Assurance Office, with consultation as required from the Centre for Teaching Excellence (CTE). Support for the technical use of the *eEvaluate* software would be provided by Information Systems and Technology (IST), and when required by the Centre for Teaching Excellence (CTE). Quality Assurance Office staff, along with the developers of *eEvaluate* (Science Computing) and IST staff, should determine an optimal strategy to ensure appropriate resourcing (sufficient capacity and operational support, user training and support) for *eEvaluate* for campus-wide use.

Quality Assurance Office staff should also monitor the performance of the course instrument and platform on a term-by-term basis, and report findings to Senate via the Senate Undergraduate Council (SUC), the Senate Graduate and Research Council (SGRC), and the Course Evaluation Advisory Group co-led by Science Computing and IST.

6.4 Data Management

Access to and use of course evaluation data are sensitive issues. As such, information generated by the course evaluations must be managed carefully, in accordance with the requirements of Policy 46 (Information Management). The collection, analysis and dissemination of course evaluation data must be carried out in accordance with best practices concerning privacy of information, transparency and accountability.

Numeric information should be made accessible after authentication by the WATIam system and should be available at the individual course level. These data should provide information about the overall course rating, course structure and instructional quality. Course evaluation data should present information to facilitate comparison with Faculty-wide ratings and program-specific ratings. Instructors should have access to all numeric information. The numeric results from course evaluations should be part of the instructor's record for annual performance review, and for tenure and promotion purposes.

Written comments from students are intended for the instructor's use only. Optional questions regarding TAs should be shared with the TAs; instructors are encouraged to engage in discussions about these results with TAs. Instructors, at their discretion, may use the written comments as the basis for seeking feedback and improvement. For example, they may show some of all of the comments to members of CTE or colleagues when seeking advice about improving their teaching or the course.

6.5 Course Evaluation Process

The project team believes that best practice for administering course evaluation includes the following obligations

1. Provide students with information about the instrument at the outset of each term in each course so that they are aware of the type of feedback that will be requested. This information should be included in the course outline
2. Orient students to the purpose and applications of the evaluation
3. Initiate the evaluation in class during the last two weeks of classes each term
4. Set aside approximately 15 minutes for in-class evaluation (for face-to-face courses); and
5. Close access to the evaluations before the start of the exam period.

7.0 Feedback Requested

In summary, the project team has identified a number of recommendations regarding the development and implementation of a new course evaluation model for Waterloo. The project team requests your feedback on the recommendations presented in the introduction of this report by answering the following questions.

In your opinion:

- | |
|---|
| <ol style="list-style-type: none">1. What are the advantages and disadvantages of Waterloo adopting a cascaded model for course evaluation?2. How well do the sample questions (see Table 1) align with the instrument design principles outlined in this report?3. What are the advantages and disadvantages of access to course evaluation information as presented in this report? |
|---|

Your responses will be used to inform the next version of this report. You can provide your feedback in the survey posted on the project's website (See: Associate Vice-President, Academic).

8.0 Next Steps

This report will be revised in response to the feedback received. The updated report will then be submitted to the Vice-President Academic and Provost. Should the Provost choose to proceed with any or all of the recommendations of this report, the recommendations would have to pass through the normal academic channels, including Senate, for approval.

If there is approval to pursue a cascaded evaluation model, a committee will be struck and user testing and survey validation should be undertaken on the core questions. Items could be developed for the additional question bank (i.e., the Faculty, academic unit, and instructor questions). The prototype instrument should be field tested through pilots, the results of which would be used to change, refine and finalize the question set.

Contact Information:

Mark Seasons, Professor
Chair, Course Evaluation Project Team (CEPT)

References

- i Ministry of Training, Colleges, and Universities. (2015). *Focus on Outcomes, Centre on*
- ii Wright, A.W., Hamilton, B., Mighty, J., Scott, J., & Muirhead, B. (2014). *The Ontario Universities' Teaching Evaluation Toolkit: Feasibility Study*. University of Windsor: Centre for Teaching and Learning. 37
- iii Chickering, A.W. & Gamson, Z.F. (1987). Seven Principles for Good Practice in Undergraduate Education. *AAHE Bulletin*.; Feldman, K.A. (2007). Identifying exemplary teachers and teaching: Evidence from student ratings. In R.P. Perry and J.C. Smart (eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective*, pp.93-143. Springer Online.; Gravestock, P. & Gregor-Greenleaf, E. (2008). *Student Course Evaluations: Research, Models and Trends*. Toronto: Higher Education Quality Council of Ontario; Hativa, N., Barak, R., & Simhi, E. (2001). Exemplary university teachers: Knowledge and beliefs regarding effective teaching dimensions and strategies. *The Journal of Higher Education*, 72, 699-729.; National Survey of Student Engagement (NSSE). (2014). *From Benchmarks to Engagement Indicators and High Impact Practices*. <http://nsse.iub.edu/pdf/Benchmarks%20to%20Indicators.pdf>; Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the validity of student evaluation of teaching: The state of the art. *Review of Educational Research*, 83, 598-642; *Undergraduate degree level expectations for the University of Waterloo*. Retrieved from <https://uwaterloo.ca/centre-for-teaching-excellence/teaching-resources/curriculum-development-and-renewal/program-review-accreditation/8-degree-expectations>; Young, S. & Shaw, D.G. (1999). Profiles of effective college and university teachers. *The Journal of Higher Education*, 70, 670-686.
- iv Weimer, M. (2010). *Inspired college teaching: A career-long resource for professional growth*. San Francisco, CA: Jossey-Bass. 75
- v McGill University. (2014). *Recommended pool of questions for courses and instructors*. Retrieved from the world wide web on March 11, 2015: http://www.mcgill.ca/tls/files/tls/recommended_pool_of_questions_2014-final.pdf; *Student Evaluation of Educational Quality (SEEQ)* – example used was employed at the University of Manitoba. Retrieved from http://intranet.umanitoba.ca/academic_support/catl/media/seeq.pdf; University of British Columbia. (2008). *University questions*. Retrieved from <http://teacheval.ubc.ca/introduction/university-questions/>; University of Toronto. (2013). *Course evaluation question bank*. Retrieved from <http://www.teaching.utoronto.ca/Assets/Teaching+Digital+Assets/CTSI+1/CTSI+Digital+Assets/PDFs/qb-nov13.pdf>; Werhun, C. & Rolheiser, C. (2014, June). *Students' motivation for participation in course evaluations outside of the classroom: The role of the instructor in online course evaluations*. Presented at the annual Society for Teaching and Learning in Higher Education conference, Queen's University, Kingston, ON.
- vi Gravestock, P. & Gregor-Greenleaf, E. (2008). *Student Course Evaluations: Research, Models and Trends*. Toronto: Higher Education Quality Council of Ontario; Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the validity of student evaluation of teaching: The state of the

art. *Review of Educational Research*, 83, 598-642.

vii Marks, P. (2012). *Silent partners: Student course evaluations and the construction of pedagogical worlds*. *Canadian Journal for Studies in Discourse and Writing*, 24(1), 1-32. Retrieved from <http://www.cjsdw.com/index.php/cjsdw/article/view/16/8>

viii Gravestock, P. & Gregor-Greenleaf, E. (2008). *Student Course Evaluations: Research, Models and Trends*. Toronto: Higher Education Quality Council of Ontario.

ix Gravestock, P. & Gregor-Greenleaf, E. (2008). *Student Course Evaluations: Research, Models and Trends*. Toronto: Higher Education Quality Council of Ontario.

x Kulik, J.A. (2001). Student ratings: Validity, utility, and controversy. In M. Theall, P.C. Abrami, & L.A. Mets (Eds.), *The student ratings debate: Are they valid? How can we best use them?* [Special issue]. *New Directions for Institutional Research*, 109, 9-25.; Gravestock, P. & Gregor-Greenleaf, E. (2008). *Student Course Evaluations: Research, Models and Trends*. Toronto: Higher Education Quality Council of Ontario.; MacNell, L., Driscoll, A., & Hunt, A.N. (2014). What's in a name: Exposing gender bias in student ratings of teaching. *Innovative Higher Education*. DOI: 10.1007/s10755-014-9313-4; Ory, J.C., & Ryan, K. (2001). How do student ratings measure up to a new validity framework? In M. Theall, P.C. Abrami, & L.A. Mets (Eds.), *The student ratings debate: Are they valid? How can we best use them?* [Special issue]. *New Directions for Institutional Research*, 109, 27-44.; Sinclair, L. & Kunda, Z. (2000). Motivated stereotyping of women: She's fine if she praised me but incompetent if she criticized me. *Personality and Social Psychology Bulletin*, 26, 1329-1342.; Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the validity of student evaluation of teaching: The state of the art. *Review of Educational Research*, 83, 598-642.; Theall, M., & Franklin, J. (2001). Looking for bias in all the wrong places: A search for truth or a witch hunt in student ratings of instruction? In M. Theall, P.C. Abrami, & L.A. Mets (Eds.), *The student ratings debate: Are they valid? How can we best use them?* [Special issue]. *New Directions for Institutional Research*, 109, 45-56.