

Using the New Course Experience Survey to Assess and Improve Teaching at UVic

Teresa Dawson and Marty Wall, Learning and Teaching Centre
February 6th, 2009

TABLE OF CONTENTS

1. Introduction
2. Best practices for administering the CES and for obtaining higher quality feedback
3. The structure of the new UVic CES
4. Analyzing the results
 - a. The summary report form
 - b. Analyzing the CES quantitative data for a specific instructor
 - c. Analyzing the qualitative data for a specific instructor
 - d. Putting the quantitative CES data in context
 - i. Use of Annual Teaching Activity Reports and development of rubrics for decisions concerning annual salary adjustments (merit pay)
 - ii. Use of the Teaching Dossier at key career moments
 - e. Taking account of issues concerning reliability and validity of the CES ratings
 - i. Reliability
 - ii. Validity
 - f. Taking account of factors that might have an impact on CES ratings: a brief review of the literature
 - i. *Level of the course*
 - ii. *Required courses*
 - iii. *Discipline*
5. Using the CES to mentor the instructor
 - a. Setting a context for formative feedback
 - b. Overall approach
6. Vision for the use of the CES at UVic

About the authors

Annotated Bibliography

APPENDIX A: Sample CES form

APPENDIX B: Sample Course Experience Survey Summary Results for Fall 2008

APPENDIX C: Example of the development of a merit assessment rubric for a department

APPENDIX D: Selected resources available for instructors linked to individual CES items

Using the New Course Experience Survey to Assess and Improve Teaching at UVic

Updated January 31st 2009

Teresa Dawson and Marty Wall

1. Introduction

The Course Experience Survey (CES) is an instrument for obtaining from every undergraduate course each term the ratings by students of the instructor's teaching and of their experience in the course through both quantitatively scaled multiple-choice items and open-ended questions.¹² The use of this uniform instrument not only allows individual feedback to each instructor but allows for some degree of comparison across instructors, courses, departments and faculties. In addition, the instrument is sufficiently flexible to provide the opportunity for departments or individuals to add items that might be of particular or distinctive use.³

It is important to point out that the CES results are not themselves direct measures of the quality of teaching and course experience. Rather the results are direct measures only of *student ratings* of the quality of teaching and course experience. There are actually a variety of ways to evaluate the quality of teaching and to garner feedback to improve it. These include peer review of teaching; assessment of syllabus and other course material; student assignment and assessment procedures; innovative elements; and the instructor's self reflections. As such, the instructor's teaching dossier (which integrates these materials from a variety of sources for key career moments such as tenure) serves as a very effective lens for assessing and improving teaching, as it provides a means for triangulating evidence (finding consistency among various independent sources of data, including the CES results) regarding the quality of teaching. Thus the information specifically garnered from the CES—student ratings—should constitute only one component (albeit an important one) of any teaching assessment process (whether formative or summative⁴) such as the Annual Teaching Activity Report⁵ or the more comprehensive teaching dossier,

The CES can serve as a useful component in assessing teaching and the student experience at various levels. For the instructor it provides direct responses from students that can serve as the basis for ongoing reflection, modification, and experimentation, thereby laying the

¹ In this sense it is designed in part to support the UVic Policy outlined in the Framework Agreement (13.8 Student Evaluations of Teaching Effectiveness) that requires (13.8.1) "Evidence of teaching effectiveness shall include complete aggregated statistical results of all teaching evaluation questionnaires administered during the period of review, in accordance with the Evaluation Policy of the Faculty in which the Faculty Member holds an appointment, or the Faculty in which the course is offered."

² Its use in graduate courses is also recommended although in this context small sample sizes may weaken the reliability of results.

³ However, such items should be designed with the same rigor and care as the original items. "Homegrown items" are usually invalid (Everly and Aleamoni, 1972) and are therefore not recommended for personnel decision-making unless they have also been tested for validity and reliability. For assistance with designing questions to address specific educational goals, as well as referral to a central database of supplemental questions used by faculties and departments across UVic, please contact the Learning and Teaching Centre, lrc@uvic.ca or 721-8571.

⁴ Summative assessment is often referred to as "evaluation" and is for personnel decision-making such as merit pay decisions, whereas formative assessment is for the purpose of teaching improvement and may be conducted by the institution but also by individual instructors interested in reflecting on their own practice.

⁵ Sometimes unfortunately somewhat misleadingly referred to at UVic as the "short-form teaching dossier." See <http://www.lrc.uvic.ca/servicesprograms/teachingdossiers.php> for more details.

groundwork for the Scholarship of Teaching and Learning.⁶ For the department, it provides a basis for examining student responses to teaching across courses, instructors, and years, both currently and longitudinally. At the institutional level, it provides norms and baselines that allow for a critical examination of teaching across faculties over time and helps to answer questions about teaching and learning at UVic and with respect to the national context (e.g. National Survey of Student Engagement).

The interaction among these various levels can be highly productive. For example, the availability of aggregate information can provide the individual instructor with useful baseline information for examining his or her own CES data relative to established baselines. The information can also provide the necessary institutional norms to support university wide nominations for national teaching awards, such as the 3M National Fellowships. *It is noteworthy that, to our knowledge, until this recent initiative UVic was the only university in its class in Canada that did not employ a uniform, university-wide instrument for the assessment of teaching by students.*

2. Best practices for administering the CES and for obtaining higher quality feedback

Consistency in the administration of the CES form to students in the class is important in order to produce fair outcomes in terms of results. Best practices for administration of the forms suggest the following:

- The instructor should not be present while students are completing the forms.
- It is advisable not to administer forms directly after a test.
- Students should be allowed 10-15 minutes to complete the forms, preferably at the beginning of the class (this signals its importance).
- One student in the class should be asked to collect the completed forms. The student must then place the forms in the envelope provided and return it to the Departmental/Faculty Office.

Best practices that help to increase the amount of useful formative information given by students (the qualitative feedback) include:

- Taking a few minutes at the end of the previous class to remind students that they will be evaluating the course next time they meet and asking them to think in the meantime about the feedback they can give to help improve it
- Setting the context—explaining the purpose of the CES at UVic
- Expressing clearly that you value their opinion and indicating how you will use their input (to improve the class next time around)

3. The structure of the new UVic CES

The UVic CES (Appendix A) contains five parts: Part I provides student ratings of the instructor's teaching. Part II provides student ratings of the course design. Part III provides ancillary information about the student's course experience, Part IV provides student ratings of optional additional matters provided by the instructor or department, and Part V allows for open-ended comments by students that might expand on their ratings to particular items in the previous parts. As opposed to Part V, responses to Parts I, II, III, and IV can be

⁶ The application of disciplinary epistemologies and scholarly research methods to the study of student learning in one's own field.

automatically scanned and analyzed quantitatively. Note that the scale points from Parts I and II range from 1 (very poor) to 5 (excellent) to allow for maximum differentiation of responses prior to the averaging of the results.

4. Analyzing the results

a. The summary report form

A preliminary examination of the sample summary report form (Appendix B) reveals some useful information. You can see from the report, for example, that there is a *response rate* of 52.4%. In this case there were 22 forms scanned out of 42 students enrolled in the course. This means that analysis of the data is based on 22 out of 42, or 52.4% of the students. These numbers have bearing on the reliability and validity of the summary information. For example a small number, say 30 or fewer, might not provide a suitable basis for a quantitative analysis. Similarly, a small response rate percentage, say 40% or fewer might raise the issue of bias based on the self selection of those students who respond. This issue would be particularly important when making judgments about an instructor for tenure and promotion. In such cases, of course, other sources of data in the Teaching Dossier would need to be used to contextualize the information.

Referring to the CES form itself (Appendix A), note that for each of statements 1 to 8 on Part I (instructor's teaching), statements 9 to 15 on Part II (course design) and statements 21 to 25 on Part IV (optional, additional), the numerals across the five scale ratings (1 to 5) represent the percentage (not the number) of students selecting each scale rating. These percentages should in each case sum to 100, barring rounding errors. In addition, the total number of students responding is provided together with the mean, median and standard deviation of the student ratings for each statement.

On the sample summary report form (Appendix B), the mean rating for each statement is defined as the sum of the individual student ratings divided by the number of them. For example, on Part I, Statement 1, the mean rating is 4.41 on the scale from 1 to 5. The standard deviation for each statement provides a measure of dispersion (*i.e.* lack of agreement) among the individual student ratings for that statement. For example, if on a particular statement the students all selected the same rating, there would be total agreement among the students, and the standard deviation for that statement would be zero.

The median, an alternative to the mean as a measure of central tendency, is defined as the middle student rating, *i.e.* the rating value for which half of the student ratings are higher in value and half are lower in value. In this example, the median rating is 5, indicating that more than half the ratings on Statement 1 were 5.

Finally, for each of statements 16 to 20 on Part III, the numerals represent the number (not percentage) of students selecting each alternative.

Results from the open-ended questions on Part V are not included in the summary report form.

b. Analyzing the CES quantitative data for a specific instructor

- i. The statements on the CES form have been repeated (slightly abbreviated in some cases) on the summary report form (Appendix B). It is often helpful to begin by noting the pattern of mean ratings across the statements. Is there a predominant rating? If there are fluctuations or anomalies, these might be worth further exploration. In such cases, the standard deviation can provide additional information not revealed by the mean, as can viewing the pattern of ratings as a whole. For example, if for a particular statement half the students gave a very low rating and the remaining half gave a very high rating, yielding a high standard deviation, this result would warrant consideration of the underlying characteristics of the students in the class.⁷
- ii. The mean student rating on Statement 8, the overall rating of the instructor, provides a global, though crude, *very preliminary* basis for comparing instructors within a unit, and the mean rating is likely to correlate with those of at least several of the other statements.⁸ However, a simple rank ordering among instructors on this basis alone is not appropriate, and a number of other factors must be taken into account (please see below, for factors that impact ratings). On its own, a comparison of the rating for this statement can provide a way of differentiating among those at the very bottom (between 1 and 2) and at the very top (between 4 and 5), to be singled out for further individual attention. It is likely that the majority of instructors will fall within 2 standard deviations of the means for their department or faculty. Those at the extremes might possibly be considered, with additional information, as candidates for more intensive mentoring or for teaching awards.
- iii. There is a clustering of statements around features of the instructor's teaching. Among the statements of Part I (Statements 1 – 7), Statements 1 – 3 form a cluster of statements pertaining to the instructor's presentation skills (preparedness, clarity, motivational capability), while Statements 4 – 7 form a cluster of statements pertaining to the instructor's attitude and helpfulness. While one might expect some degree of correlation among the mean ratings within (and perhaps between) these two clusters, one might pay particular attention to cases where there appear to be anomalies or where the mean ratings are quite different from those of Statement 8, the overall rating.
- iv. Similarly, the mean student rating on Statement 15, is the overall rating of the course as an effective learning experience. It provides the “global” counterpart for the course design that Statement 8 represents for the instructor's teaching. One would expect a congruency in ratings between Statements 8 and 15. The lack of such congruency would be worthy of further exploration. Sometimes this is explained simply by the fact that the teacher of the course is not the one who designed it.

⁷ Such circumstances are often ones where an individual consultation with the Learning and Teaching Centre can be of assistance in teasing out particularly complicated underlying issues—for example where there is a bi-modal distribution in the ratings. For an appointment please contact lrc@uvic.ca or 250 721-8571.

⁸ Research will be required at UVic, once we have gathered sufficient data, to determine to what extent each statement correlates with the overall rating.

- v. As in the case of the clustering of statements around features of the instructor's teaching (iii, above), there is a clustering of statements around features of the course design. Among the statements of Part II (Statements 9-14), Statements 9 – 12 form a cluster of statements pertaining to various features of the course design itself, while Statements 13 and 14 pertain to assessment and external relevance. Again, while you might expect some degree of correlation among the mean ratings within (and perhaps between) these two clusters, you might pay particular attention to cases where there appear to be anomalies; where the mean ratings are quite different from those of Statement 15, the overall rating; and/or where the standard deviation is high, suggesting a significant lack of agreement among the students. Moreover, apart from their use in refining the interpretation of CES ratings on the instructors' teaching, student ratings on Part II can also be useful in course redesign and curricular planning.⁹

c. Analyzing the qualitative data for a specific instructor

According to the Faculty Agreement (13.8.3 Student Comments) "A Faculty Member may choose to include anecdotal or subjective comments by students or former students in her or his teaching dossier."¹⁰ Where such comments are included that have been collected as part of a survey of students in a course, all the comments from that course must be included in the Faculty Member's teaching dossier." This latter stipulation is there to prevent selection of only the favorable comments on the part of the instructor. However, unfortunately this rule has produced the perhaps unintended outcome at UVic that many instructors do not regularly include their comments as part of their annual activity reports or teaching dossier because they may be concerned that their chair will focus on only one or two outliers contained therein. This is extremely unfortunate since it prevents a large amount of very valuable data from being available to help inform mentoring of the instructor.

The approach to interpreting the qualitative data should be somewhat different from that of the quantitative data in that the focus should be more on formative, rather than summative, feedback. Some of this information can be highly informative, but it should be noted that not all of the students write comments, and the responses themselves could be affected to some extent by self selection of students who choose to respond and by ambiguity and capriciousness in some of the responses. The appropriate strategy is to look for patterns of responses across students and for the seeming congruency or lack of it between these responses and the quantitative data. While all responses are potentially important, one should beware of overemphasis on outlier responses from individual students that vary significantly from the general pattern. One of most effective ways to use the qualitative data is during consultation of the instructor with a mentor or with the chair, since student written feedback can be a particularly fruitful source of positive suggestions for improvement.

⁹ The LTC provides a tiered series of support services for curricular planning including: co-hosting a workshop for chairs and deans on facilitating complete departmental program review (often in response to external academic review); supporting departmental retreats for program redesign; offering Course Redesign Workshops (particularly for teams designing foundation or capstone courses); and providing one-on-one consultations for faculty on individual courses. For information please see <http://www.ltc.uvic.ca/> or call (250) 721-8571.

¹⁰ The authors would prefer the term "qualitative data" since "anecdotal" implies they are somehow less legitimate.

d. Putting the quantitative CES data in context

i. Use of Annual Teaching Activity Reports and development of rubrics for decisions concerning annual salary adjustments (merit pay)

The CES data provide one set of indices about an instructor's teaching, but other independent forms of data are, of course, necessary to provide a rounded picture of the instructor's teaching profile. As such, the departmental/faculty Annual Teaching Activity Reports (sometimes somewhat misleadingly referred to at UVic as "short-form teaching dossiers") encourage the provision of important contextual information from a variety of sources such as peer reviews,¹¹ student comments, and other teaching dossier-type data that can reinforce, explain, and amplify the CES data.¹²

Best practice already established at UVic in many departments/faculties is that the annual salary adjustment (merit pay) committee meets a year ahead to establish transparent criteria for that year and that, at this time, consistent formulae and assessment methods are agreed upon and published together with these criteria, so that there is clarity regarding expectations for the coming year.¹³ One example of how this might unfold is illustrated in Appendix C. However, the authors would argue that the exact form of the final outcome of this process (provided it is seen as fair by colleagues) is far less important than the *process* of its production. The process of a department or faculty coming together to debate definitions of excellence in teaching in their discipline(s) and the relationship of various instructional indicators to departmental/faculty strategic planning goals for the curriculum in their area (when facilitated with care) has a value that is hard to overestimate.

ii. Use of the Teaching Dossier at key career moments

At key career moments—such as tenure, promotion, and/or submission for a major career teaching award—the use of contextualizing information is even more important and takes the form of a full teaching dossier.¹⁴

e. Taking account of issues concerning reliability and validity of the CES ratings

i. Reliability

A measure is said to be reliable if it tends to give a similar result when repeated. CES ratings are reliable to the extent that the results are repeatable under similar

¹¹ One suggested approach to carrying out peer reviews is described on the LTC web site at <http://www.ltc.uvic.ca/servicesprograms/teachassess.php>

¹² See <http://www.ltc.uvic.ca/servicesprograms/teachingdossiers.php> for more information and templates for Annual Teaching Activity Reports for faculties.

¹³ The LTC welcomes the opportunity to collect examples of best practice at UVic and to share those with others by request.

¹⁴ For LTC services that support the development of Teaching Dossiers please see <http://www.ltc.uvic.ca/servicesprograms/teachingdossiers.php>.

circumstances. Factors affecting CES reliability include, for example, class size (classes under ten threaten reliability) and short-term influences (giving a very hard or very easy test just prior to the CES exercise). Under normal circumstances, well designed CES instruments are considered reliable (Aleamoni, 1999).

ii. Validity

A measure is said to be valid if it is indeed measuring what it intends to measure (*e.g.* important aspects of instruction). Reliability is a necessary but not sufficient condition of validity. CES ratings are valid to the extent that the results correlate with other measures of teaching effectiveness. For example, student IQ might be a *reliable* measure of the instructor's teaching effectiveness (the result is repeatable), but it is not *valid*, since it does not correlate at all with other measures of teaching quality. Well designed CES instruments are considered to be valid in that they correlate positively with other measures of teaching effectiveness (Aleamoni, 1999).

We have every reason to anticipate that the new UVic form will be both valid and reliable under normal circumstances for three reasons¹⁵:

- It is based on well-established psychometric principles
- Questions are drawn largely from questions used in other institutions that have been well-tested (20+ years of data) and found to be both reliable and valid
- Questions were selected specifically to address UVic's carefully defined educational goals

In contrast "homegrown" forms are likely to have been constructed by committees and do not usually employ psychometric expertise. They are therefore likely to be neither valid nor reliable (opening up considerable possibility for bias).

f. Taking account of factors that might have an impact on CES ratings: a brief review of the literature

There is an enormous literature on issues concerning external factors that might affect student ratings of instructors (Arreola, 2007).¹⁶ Some are potentially problematic but can be dealt with effectively by using reliable and valid CES instruments under careful conditions. Others have to be seriously considered in judging teaching effectiveness using the CES results.

According to the research, *assuming a valid and reliable instrument*, the following **have been shown to impact ratings**:

- *Level of the course*—students in first and second year do tend to rate instructors lower than students in upper level classes. This is important because it used to be thought that it was class size that mattered (that larger classes rated lower than smaller ones). It is now established that this is not the case,

¹⁵ This said, we will need at least a term's worth of data at UVic to be able to ascertain for sure.

¹⁶ For an excellent, comprehensive review of the research addressing these factors, see also Aleamoni, 1999.

and it is the level of the course (which of course is often conflated with class size) that counts (Arreola, 2007).

- *Required courses*—students in required courses tend to rate instructors lower than students in non-required ones.
- *Faculty*—mathematics and science courses generally receive lower ratings than those in social sciences or humanities. It should be noted that this is not because the former have poorer teachers but rather that there is a faculty effect. If the authors' previous experience in other institutions holds true, with time we should be able to predict the faculty norms for our UVic ratings. Such norms, once established, tend to be stable over time.

Again, according to the research, *assuming a valid and reliable instrument*, the following, **have NOT been shown to impact ratings**:

- Time of day the course is taught—early morning classes or after lunch classes do not automatically get lower ratings
- Popularity (the issue of student ratings as merely representing a “popularity contest”)
- Giving disproportionately high (or low) grades—you cannot “buy” good ratings
- Class size (this is a **new** finding)—large classes do not automatically get lower ratings, *but please see below re level of the course*
- Gender of the instructor—women do not get consistently lower ratings than men

5. Using the CES to mentor the instructor

a. Setting a context for formative feedback

Perhaps the most important goal of the CES is to enable it to support a process of ongoing teaching improvement in the department/faculty (and across the university). As such, the chair or dean plays a vital role in this process and the CES should be seen in this context. For the instructors, the CES data become part of their record of annual activity but, perhaps more importantly, should be used longitudinally as part of the evolution of the instructor's teaching dossier (or portfolio) to explore individual instructional themes, innovations and experiments¹⁷; document enhancements; support claims to excellence in particular areas; and document pedagogical career development in a scholarly manner. In this regard, the CES can serve an interactive role, stimulating the chair or dean to make a connection with the individual instructor concerning teaching and learning in the department.

If the chair (dean) is meeting with an individual instructor, it is advisable for the chair (dean) and the instructor to have both seen the summary report results ahead of time. The authors of this manual also find the individual written comments from students extremely helpful in elucidating the quantitative information and highly recommend their use but

¹⁷ In this sense, “dips” in ratings initially for one or two items may result from an innovation or pedagogical experiment and are not necessarily negative (as in any scholarly endeavor experimentation may not succeed the first time around), hence the need to allow the instructor to contextualize their work.

realize that this may not be possible depending on the nature of individual/faculty departmental policies and environment for discussion of teaching and learning matters.¹⁸

b. Overall approach

As a general principle the authors would advocate taking a scholarly approach to any analysis of CES results (and of course to the assessment of teaching in general). A useful rule is to approach this analysis just as you would any statistical data in your own research. This means not generalizing from an overly small sample size, or focusing only on single outliers.¹⁹ It is helpful to set a precedent in the department or faculty that, for example, you meet each year with everyone in their first three years, or whatever seems appropriate. In this way you do not single out particular individuals but rather establish a culture of mentoring. This means that, if an issue were to arise, this would not be the circumstance under which you would first see the instructor to talk about teaching.

Before the meeting, it is helpful to ask to review a copy of one of the instructor's syllabi, a brief statement of his/her approach to teaching, and a sample test/assignment/other form of assessment. In this way you should be able to see if his/her pedagogical goals, as stated, line up clearly with the learning goals of the course (evidenced on the syllabus) and with the assessment methods being used. This analysis will allow you to discuss methods and approaches that are likely to work well for the individual instructor, even if he/she has different pedagogical goals from your own.

You also perform the preliminary analytical process described in 4b) above. The focus here is less on relative comparisons of the instructor's ratings with the ratings of other instructors than it is on a discussion of the pattern of the mean ratings and of cases of particular lows or highs in the ratings for this instructor.

The beginning of the discussion could be then informed by the agreement, or lack thereof, of these ratings with

- the instructor's own articulated views and aspirations about the learning outcomes for the course and where these do or do not conform with the student ratings
- comparisons with CES ratings in courses the instructor has taught previously, if applicable
- comparisons with CES ratings in other courses currently taught by the instructor
- the student comments in Part V, if available

¹⁸ Please note that the process outlined here is only intended as a suggestion and will vary considerably from department to department. In addition, if a chair/dean is new to the position or new to UVic and has not had occasion to have experience in analyzing and providing formative feedback on CES-type results before, he/she might benefit from a consultation with the Learning and Teaching Centre before conducting such meetings. The LTC has on staff faculty who have considerable collective experience (including as a chair) in conducting such formative discussions. Similarly, if for some reason, a chair or dean is not comfortable conducting such discussions, or if individual instructors wish for further assistance, the LTC can serve this role. All consultations are considered confidential, and it is always most productive for all if the individual does not feel they have been "sent" for consultation but rather has chosen to ask for it.

¹⁹ The only exception to this might be if you see concerns of bias (very occasionally you might see a student concerned about classroom climate issues such as racism) in which case this can be worth exploring (with extreme care and without any assumptions in either direction of course) because perhaps only one student felt sufficiently comfortable to raise such a difficult issue.

The subsequent discussion can then focus on areas of the instructor's strengths and where certain areas would benefit from some changes. Attention would be paid to past and future teaching and how the instructor is developing as a teacher. The chair should feel free, when appropriate, to provide positively oriented referrals to UVic teaching resources for individual consultation and workshops, for example, on issues of individual teaching development, teaching grants, and the preparation of the Teaching Dossier. For further information on such resources, linked specifically to the questions on the new CES see Appendix D and the LTC website: <http://www.ltc.uvic.ca/>

6. Vision and goals for the CES at UVic²⁰

In close collaboration with our colleagues in Institutional Planning and Analysis and in Computing and Systems Services, as well as with those in the departments and faculties, the Learning and Teaching Centre suggests the following vision and goals for the development of the CES program at UVic.

- a. Assisting colleagues in using the CES to improve teaching and learning across the campus and beyond and hence to enhance the student experience (formative uses).
- b. Supporting the institution in its summative needs for evaluation (for merit pay, tenure and promotion) in a manner that is fair, transparent, supportive, collegial and in line with the formative goals (above) as well as the Faculty Framework Agreement and other policies of the university.
- c. Supporting faculty colleagues in effectively documenting their teaching and learning work as part of their academic work. This includes advocating for the triangulation of data from a variety of sources through the Teaching Dossier in order to contextualize the CES ratings appropriately.
- d. Facilitating chairs and deans in implementing curriculum design and redesign in their departments and faculties by allowing them to see patterns in student responses from the "course design" aspects of the report.
- e. Enhancing our ability to apply for regional and national awards and to make informed decisions about on-campus awards, as well as to be proud of our teachers in the national context more generally.

This in turn requires

- An integrated database with appropriate security features capable of dealing easily with longitudinal research and query functions, as well as the ability to deliver reports in a secure environment. For example, the establishment of longitudinal norms makes it possible to track developments in the student response to teaching at UVic at various levels. At the university level, the presence of norms across the faculties allows for fair assessment of individual instructors for university level awards such as the Harry Hickman Alumnae Award. At the national level, the presence of university norms allows for the presentation of reliable and valid comparative assessment measures of teaching to support nominations for national awards such as the 3M National Teaching Award.
- Employment of well-established psychometric principles in survey design as well as drawing on the most recent (and extensive) research in the field

²⁰ As presented to Deans Council on February 3rd 2009

- Developing a careful research design and implementation process drawing on the professional and academic expertise that exists at UVic
- Checking for validity and reliability of the tools; establishing norms for faculties and for the factors that are known to affect ratings within faculties.
- Developing support materials at all levels (chairs, deans, individual instructors) and communicating clearly to all members of the campus community

Annotated Bibliography

Aleamoni, L.M., 1999. Student rating myths versus research facts from 1924 to 1998. *Journal of Personnel Evaluation in Education*, vol. 13, No. 2, pp. 153-166.

A comprehensive review of the student ratings literature. Extraordinarily useful and well-researched summary.

Arreola, R.A., 2007. *Developing a Comprehensive Faculty Evaluation System: A Guide to Designing, Building, and Operating Large-Scale Faculty Evaluation Systems*. Third Edition. Bolton MA: Anker Publishing Company Inc.

Covers everything from the start of the process to implementation. Includes banks of questions that have been tested linked to a large variety of learning outcomes. Arreola is highly regarded for his long time work in this area. He has helped a large number of institutions in North America with his expertise. Chapter 12 is a favourite since it deals with all the “myths” regarding factors that do or do not impact evaluations.

Boyer, E. (1990). *Scholarship reconsidered: Priorities of the professoriate*. San Francisco, CA: Jossey-Bass.

This is now a classic. Boyer offers a new paradigm that recognizes the full range of scholarly activity by college and university faculty (which he divides into the scholarships of discovery, integration, application and teaching and learning), and questions the existence of a reward system that pushes faculty toward research and publication and away from teaching and integrating and applying existing knowledge.

Centra, J.A., 1977. Student ratings of instruction and their relationship to student learning. *American Educational Research Journal* Vol. 14, No. 1, pp. 17-24.

Still an old favorite. The title says it all. Indicates that student ratings, if well constructed, can measure learning outcomes.

Everly, J.C., and L.M. Aleamoni, 1972. The rise and fall of the advisor ... students attempt to evaluate their instructors. *Journal of the National Association of Colleges and Teachers of Agriculture*, Vol. 16, No. 2, pp. 43-45.

Deals with the particular issue of the lack of validity and reliability in “homegrown” instruments, which constitute the majority of current forms still used in North America.

Glassick, C. E., Huber, M. T., & Maeroff, G. I. (1997). *Scholarship assessed: Evaluation of the*

professoriate. San Francisco, CA: Jossey-Bass.

Examines the changing nature of scholarship in today's colleges and universities and proposes new standards for assessing Boyer's four types of scholarship (see above) and evaluating faculty, with special emphasis on methods for documenting the effective scholarship of teaching and learning.

Huber, M. T., & Hutchings, P. (2005). *The advancement of learning: Building the teaching commons*. San Francisco, CA: Jossey-Bass.

Explores the premise that the scholarship of teaching and learning holds the key to improving the quality of learning in higher education. The authors offer illustrative examples of the scholarship of teaching and learning, explore what it takes for faculty to become engaged in such work, and detail the structures and policies that campuses need to develop and support faculty engagement with the teaching commons.

About the authors

Teresa Dawson is Director of the Learning and Teaching Centre at UVic. Previously, in similar roles, she has been charged with implementing or extensively refining university-wide course evaluation systems at both the University of California—Los Angeles (UCLA) and the University of Toronto—Scarborough Campus (UTSC). Teresa has spent much of her career advocating for the effective and fair assessment of teaching using teaching dossiers to, amongst other goals, contextualize course experience survey-type data. She was the author of the “UTSC Guidelines: Assessment of Teaching Effectiveness for Decisions Concerning Tenure, Promotion to Senior Lecturer and Promotion to Full Professor” (passed by its Council in 2004).

Marty Wall served as advisor to the taskforce on the design and implementation of the CES at UVic from 2007-08. Previously, Marty was Chair of Psychology at the University of Toronto 1991-2001 where he introduced innovative rubrics and processes for the assessment of teaching for the purposes of merit pay, tenure and promotion, as well as for the improvement of teaching and the mentoring of junior colleagues. At U of T Marty was also responsible for chairing the taskforce that designed and implemented the U of T Faculty of Arts and Science course evaluation survey in the mid 1990s that is now used extensively throughout the university including at UTSC. Marty has won numerous student awards for his teaching and is a national 3M Teaching Fellow.

APPENDIX A: Sample CES form



University of Victoria

Course Experience Survey

Course Name: _____ Course Section#: _____ Instructor: _____
 (e.g. HA 260) (e.g. Y01)

The University of Victoria values your feedback on your course experience. Your responses will be considered in the assessment of instructors and for improving the design of courses. All your answers are anonymous. Please contribute your ratings by responding to the questions below. Thank you!

I Instructor's teaching.

Shade circles like this:

Please provide your rating on the following statements:

	Very Poor 1	Poor 2	Adequate 3	Good 4	Excellent 5
1. The instructor was prepared for course sessions.....	<input type="radio"/>				
2. The instructor's explanations of concepts were clear.....	<input type="radio"/>				
3. The instructor motivated you to learn in this course.....	<input type="radio"/>				
4. The instructor was available to answer your questions or provide extra assistance as required.....	<input type="radio"/>				
5. The instructor ensured that your assignments and tests were returned within a reasonable time.....	<input type="radio"/>				
6. The instructor was helpful in providing feedback to you to improve your learning in this course.....	<input type="radio"/>				
7. The instructor demonstrated respect for students and their ideas.....	<input type="radio"/>				
8. Overall, the instructor was effective in this course.....	<input type="radio"/>				

II Course design.

Please provide your rating on the following statements:

	Very Poor 1	Poor 2	Adequate 3	Good 4	Excellent 5
9. The course structure, goals and requirements were clear.....	<input type="radio"/>				
10. The materials provided for learning the course content (e.g. handouts, posted material, lab manuals) were clear.....	<input type="radio"/>				
11. The assigned work helped your understanding of the course content....	<input type="radio"/>				
12. The course provided opportunities for you to become engaged with the course material, for example through class discussions, group work, student presentations, on-line chat, or experiential learning.....	<input type="radio"/>				
13. The methods of assessment used to evaluate your learning in the course were fair.....	<input type="radio"/>				
14. The course provided relevant skills and information (e.g. to other courses, your future career, or other contexts).....	<input type="radio"/>				
15. Overall, the course offered an effective learning experience.....	<input type="radio"/>				

(over...)

III Statements about yourself.

16. My primary reason for taking the course:
- | | | | | |
|-----------------------|-----------------------|--------------------------|-----------------------|-----------------------|
| Interest | Program requirement | Reputation of Instructor | Reputation of course | Timetable fit |
| <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
17. The approximate number of classes or labs that I did **not** attend:
- | | | | | |
|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Missed fewer than 3 | Missed 3-10 | Missed 11-20 | Missed more than 20 | Not Applicable |
| <input type="radio"/> |
18. Relative to other courses I have taken at UVic, the workload in this course was:
- | | | | | |
|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Extremely heavy | Somewhat heavy | Average | Somewhat light | Extremely light |
| <input type="radio"/> |
19. The approximate number of hours per week I spent studying for this course outside of class time:
- | | | | | | |
|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Less than 1 | 1 to 2 | 3 to 5 | 6 to 8 | 9 to 10 | More than 10 |
| <input type="radio"/> |
20. As a result of my experience in this course, my interest in the material:
- | | | |
|-----------------------|-----------------------|-----------------------|
| Decreased | Stayed the same | Increased |
| <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

IV Additional questions:

- | | Very Poor
1 | Poor
2 | Adequate
3 | Good
4 | Excellent
5 |
|-----|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 21. | <input type="radio"/> |
| 22. | <input type="radio"/> |
| 23. | <input type="radio"/> |
| 24. | <input type="radio"/> |
| 25. | <input type="radio"/> |

Thank you for completing this course experience survey!

APPENDIX C

Table 1: Example of the development of a merit assessment rubric for a department based on its strategic goals (percentages vary for Senior Instructors)

Research (including scholarships of discovery, integration and application) 40%			Teaching (including scholarship of teaching & learning) 40%			Service (including contributions to community) 20%		
Component	Evidence example	Wgt	Component	Evidence example	Wgt	Component	Evidence example	Wgt
Productivity	Publications		Instructional design			Community participation	Committee service	
Impact	Citations, Legislation changes		Instructional delivery			Advocacy	Pro-bono service on NGO board	
etc...			Instructional Assessment			Integration of scholarship and community	Community based installation, Policy documents	
			Course Management			etc...		
			Undergraduate mentoring or support outside of class					
			Graduate /post doc mentoring					
			Integration of teaching and research					
			SOTL					
			etc...					
TOTAL			TOTAL			TOTAL		

Ideally the role components identified in the left hand side of each of the three columns (for research, teaching and service/community) would relate directly to the departmental/faculty strategic planning goals and hence indirectly to the University's Strategic Planning goals. For example, if the department has as a goal integrating research and teaching, or the development of a foundation or capstone course, or the accommodation of diversity amongst students, such outcomes would be evidenced in the items identified in the component columns (e.g. development of new foundational curriculum, syllabi open to diverse learning styles, high values on particular CES questions such as those relating to the approachability of the instructor, etc.)

For each of the three columns the sources of evidence would then likely be broken down further. If we focus here, however, on the (central) teaching column it might look something like the following to allow for the triangulation of data sources (adapted from Arreola, 2007).

Table 2: Expansion of central column from Table 1 above

Teaching (including scholarship of teaching & learning) X/40				
	Sources of evidence: examples			
Components	Students	Instructor	Peers and colleagues	Weighted rating
Instructional design <i>(example)</i>	CES question 15	New syllabus for course XXX	Assessment of team teaching in new course designed by colleague	
Instructional delivery <i>(example)</i>	CES question 8	PPT example of lecture Case study	Peer visit to class	
Instructional assessment				
Course Management				
Undergraduate mentoring or support outside of class				
Graduate /post doc mentoring				
Integration of teaching and research				
Scholarship of teaching and learning				
			TOTAL	

APPENDIX D: Selected resources available for instructors linked to individual CES items

Item characteristic (abbreviated)	Selected resources to help
1. Preparation for course sessions	“Generating weekly/session outlines” http://www.ltc.uvic.ca/servicesprograms/resources/generating.php
2. Explanations of concepts were clear	Instructional Skills Workshops and Presentation Skills Workshops “Teaching Tips” Series “PowerPoint as a teaching device”
3. Instructor motivated you to learn	“Student Success” booklet http://www.ltc.uvic.ca/servicesprograms/publications/documents/TERCb ookletAugust08v2-MostCurrentVersion.pdf “Teaching Tips” series “Strategies for large class teaching”
4. Availability to answer questions	UVic “Critical Incidents” DVDs and Workbooks “Essentials for Career Success” (Orientation for New Faculty – session on student expectations)
5. Assignments/tests returned in reasonable time	“Help with Grading” http://www.ltc.uvic.ca/servicesprograms/resources/index.php “Working with TAs” “Balancing your time: some tips to help you succeed”
6. Feedback improved learning	“Creating a grading key” “Assignment Design and Feedback”
7. Demonstrated respect for students	“Suggested Accessibility statement for your syllabus” Course Redesign Workshop on Diversity
8. Overall instructor effectiveness	Intensive Summer Teaching Institute http://www.ltc.uvic.ca/servicesprograms/orientation.php “Balancing your time: some tips to help you succeed”
9. Course structure, goals and requirements were clear	Course Redesign Workshop “Teaching Tips” series (“Preparing a Great Syllabus”)
10. Materials provided for learning content were clear	“Teaching Tips series (“Teaching with PowerPoint – Revisited,” “Preparing a Great Syllabus”) Teaching and Learning with Technology at UVic (newsletter/blog)
11. Assigned work helped you understand course content	“Assignment design and feedback”
12. Course provided opportunity to engage with material	Course Redesign Workshop “Teaching Tips” series (“Active Learning Strategies”)
13. Assessment methods were fair	“Help with Grading” http://www.ltc.uvic.ca/servicesprograms/resources/index.php “Assignment design and feedback” “Teaching Tips” series “Optimizing the Multiple-Choice Exam” and “Grading and evaluation” “Academic Integrity” website http://www.ltc.uvic.ca/servicesprograms/resources/integrity.php
14. Course provided relevant skills and information	Course Redesign Workshop “Teaching Tips” series “The First Class”
15. Overall effective learning experience	“Student Success” booklet http://www.ltc.uvic.ca/servicesprograms/publications/documents/TERCb ookletAugust08v2-MostCurrentVersion.pdf

- Further books, journals and resources are available at the LTC in the Judson Learning and Teaching Collection. For a list of topics, see <http://www.ltc.uvic.ca/servicesprograms/resources/resmaterials.php>.
- The course “Learning and Teaching in Higher Education” (EDCI 560) is jointly offered twice a year by the LTC and the Department of Curriculum and Instruction.
- The LTC *Teaching Tips Series* runs continually throughout the year and addresses many of these topics <http://www.ltc.uvic.ca/events/TTSeries.php>.