

Joint Committee on the Administration of the Agreement

Sub-committee on Course Evaluations (Article 7.10)

Final report, May 2014

1 INTRODUCTORY REMARKS

Under the terms of Article 7.10 of the 2012-2015 Collective Agreement between the York University Faculty Association and the York University Board of Governors, the parties created a joint sub-committee on student course evaluations. The committee met from April 2013-January 2014 “to develop a common set of questions, the results of which will be made available to students.”

Members: Markus Biehl, Simone Bohn, Paul Grayson, Tamara Kelly, Susan Murtha, Alice Pitt, Duff Waring and Diane Woody.

The joint-subcommittee has concluded its work and is pleased to provide the parties with a report that documents our process, provides a set of common questions and methods used to arrive at these, and identifies implementation issues for consideration by appropriate collegial governance bodies.

We began our work by consulting research literature on student course evaluations (see Appendix A) and reviewing the existing course evaluations in use at York University (see Appendix B for a list of Faculty course evaluations reviewed). We met with ASCP on June 12, 2013 to discuss the project.

During our early discussions, we identified several principles to guide our work:

- Common questions will evaluate at the level of the course, not the instructor, and shall avoid measures of ‘student satisfaction’.
- A pilot study is essential to ensure that questions are reliable across the range of program offerings, course formats, and delivery modes at York.
- A protocol for the administration of surveys of this type needs to be in place across the institution. In addition, the reporting of survey results needs to be sensitive to the bases for comparison and interpretation.
- Student surveys should include well-defined contextual questions that are intended to assist in the analysis and understanding of the data.

Once we had identified the most useful and pertinent constructs, we developed draft questions for each construct. We then designed a focus group study to test for validity and reliability. See

Appendix C for a description of the focus group study plan and results. Following that, using the results of the focus group study, we refined the draft questions and undertook a ‘broader ’ pre-test study (See Appendix D) in order to verify further the reliability of the questions in a sampling of actual courses. We are now confident that the questions we have developed meet a satisfactory level of reliability and validity.

2 SUMMARY OF LITERATURE CONSULTED

To develop a theoretical framework, members did background reading to gain a shared understanding of course evaluations and to become familiar with current research. Among the documents consulted were a HEQCO report (2008), a very recent monograph by Kember and Ginns on the topic (2012), a series of articles, and evaluation instruments used in various institutions, including some that had been validated through psychometric testing. Appendix A provides a list of resources consulted.

A document that figured prominently in our deliberations is a HEQCO report: *Student Course Evaluations: Research, Models and Trends* by Pamela Gravestock and Emily Gregor-Greenleaf (2008). The document, relevant for the Canadian context, provided a comprehensive summary of current research on student course evaluations (with extensive bibliography) and a succinct presentation of key issues and the findings regarding the validity and utility of student course evaluations.

Reading the document gave committee members a common knowledge base and a shared understanding of the following:

- current state of research on course evaluations within a context of the broader question of the evaluation of teaching generally;
- variation in the way students, faculty and administrators perceive course evaluations;
- range of evaluation instruments, policies and processes used in 22 post-secondary institutions in Canada and the US;
- importance of reliability, validity and interpretation of course evaluation data;
- importance of ensuring the validity and usefulness of instruments used.

In addition to the HEQCO report, members consulted the 2012 monograph *Evaluating Teaching and Learning: a practical handbook for colleges, universities and the scholarship of teaching* by Kember and Ginns. Our attention was focused on the chapters pertaining to evaluation principles and questionnaire design. We also consulted the chapter that presented a bank of items to be considered for course evaluations.

Various articles (see Appendix A for complete list) allowed members to deepen their awareness of issues that arise with course evaluations, e.g. variables that affect evaluation results, multidimensional measures of teaching and course effectiveness, and students’ capacity to evaluate teaching and learning.

Finally, members consulted a number of evaluation instruments used at other universities, including some that had been validated by psychometric testing. The list includes:

- SEEQ questionnaire (Student Evaluation of Educational Quality) used at Mt Allison and the University of Manitoba;
- IDEA questionnaire (Individual Development and Educational Assessment) used at Kansas State University;
- Evaluations used at the University of Alberta, University of British Columbia, Dalhousie University, Harvard University, McGill University, University of Michigan, University of Minnesota, Queen’s University, Ryerson University, University of Toronto, University of Victoria.

3 OVERVIEW OF POLICY GUIDELINES AND COURSE EVALUATIONS AT YORK UNIVERSITY

In addition to reviewing the literature and selected questionnaires used by other universities, the joint sub-committee reviewed existing policy and course evaluations at York University. The Senate Policy on Student Evaluation of Teaching, dated 1996/04/14, provides guidelines for the requirements and administration of course evaluations. Under current guidelines (Senate Policy Student Evaluation of Teaching, 1996/04/14), the results of course evaluations are to be available to the instructor (after submission of final marks), curriculum committees (and others) responsible for curricular review, Department/Division Chairs, and Tenure and Promotion Committees. The latter requires access to individual results and comparative data. Moreover, the Policy also states that units “may develop their own policies for further distribution of results” and “encourages Departments/Divisions/Faculties to make results available to students.”

An on-line bank of questions was developed at York University in order to facilitate the development of on-line questionnaires. A significant number of paper-based questionnaires are also in use at York University. In some cases, there is a Faculty-wide questionnaire, and in other cases departments within Faculties have developed their own. This variety reflects the broad range of disciplines, course formats and delivery modes that is on offer at York. The questionnaires vary in length from a very few questions to more than 30 questions. Most provide distinct sections for course and instructor evaluation. Most also provide space for written comments, guided or open-ended. A variety of rating scales is used. It is within this landscape of variation in existing questionnaires and the framework of current research on course evaluations (as outlined in the preceding section) that the Committee proceeded to develop core questions.

4 DEVELOPMENT OF QUESTIONS FOR COURSE EVALUATION QUESTIONNAIRE

Committee members opted to define a course as a structured learning experience and first analyzed existing evaluations and samples from the research literature to create a list of items that we originally called “buckets” but which are termed “constructs” in the research literature. Items included clarity and organization, objectives, course materials, learning activities, assessment, feedback, challenge/stimulation/motivation, learning space and technology, and student recommendation of the course. After deliberation, the Committee identified four constructs as the most pertinent: organization, course materials, learning activities and learning experience, and from these constructs formulated 10 core questions. In addition, four context questions were prepared for testing with students. See Appendix C. 6 for initial questionnaire.

The questions were first discussed with students, both undergraduate and graduate from across the Faculties, in 6 focus groups averaging 90 minutes and ~8 students per group. Responses from students revealed that students did not understand sufficiently well the meaning of the term “learning outcomes”; however, their responses indicated that they did interpret correctly the term “learning objectives” and thus that is the wording adopted in the revised version of the core questions. As well, two context questions were dropped based on student reactions. Finally, some minor revisions to wording were made to ensure clarity (see Appendix C for detailed discussion).

Subsequently, the revised core questions were pilot tested. This allowed the Committee to determine the overall reliability and validity of the questionnaire, and to identify and eliminate redundant questions (see Appendix D). The final questionnaire, containing 7 substantial questions in two constructs, plus two background questions, is shown below.

5 RECOMMENDED SURVEY

Dear Student,

York University uses a common set of course evaluation questions across all its Faculties and courses. The means of the answers to the seven course evaluation questions below will be made public.

Thank you for this important feedback!

1.1 A comprehensive roadmap of the course (i.e., a syllabus or course outline, a breakdown with topics, assignments and exams etc.) was provided.

Scale for question sets 1 and 2: Strongly agree (7)
Agree (6)
Somewhat agree (5)
Neither agree nor disagree (4)
Somewhat disagree (3)
Disagree (2)
Strongly disagree (1)

Not applicable (9)

1.2 The roadmap provided was mostly followed.

1.3 Important policies and procedures (e.g., attendance, participation, missed tests, late assignments, contacting instructors, etc.) were stated in the course outline.

2.1 The course materials (e.g., course kits, textbooks, readings, audio visual materials, lab manuals, websites, etc.) helped me achieve the course objectives.

2.2 The course activities (e.g., lectures, discussions, simulations, assignments, exercises and presentations, etc.) helped me achieve the course objectives.

2.3 The course tests/exams or final paper/essay were directly related to the course objectives.

2.4 The course helped me grow intellectually.

3.0 Was this course mandatory for you?

Yes (1)

Yes, as a choice from a list of required courses (.5)

No (0)

4.0 Approximately how many lectures/seminars/sessions of this course did you attend (physically or online)?

All of them (1)

Almost all of them (0.9)

More than half (0.7)

About half of them (0.5)

Fewer than half of them (0.25)

6 NEXT STEPS AND FUTURE CONSIDERATIONS

The joint sub-committee's mandate was "to develop a common set of questions, the results of which will be made available to students." In this section, we provide observations gleaned from our review of the research literature and aspects of our discussions that informed the work as it unfolded. Just as it became clear that psychometric testing was necessary to the development of the set of questions, it is also apparent that our current practices have not fully anticipated common questions as a sub-set of local questionnaires or the making of results available to students. Moreover, the existence of robust research on effective development of student course evaluations suggests opportunities for further discussion and development of resources and workshops for the York University community. Instructors who have access to their own raw data (minus identifying information) will possess more powerful information about the courses they design and/or teach. As a practical consideration, it will also be necessary for Faculties to review their existing questionnaires with a view to eliminating those questions that overlap with the set of common questions. It is also recommended that the common set precede faculty-based questions to prevent question order effects from taking place. The proposed order would invite students to reflect first on the course as a structured learning experience before reflecting on instruction and other aspects of the learning situation.

Our review of the literature suggests the need to address the following issues:

- Provision of information to students describing processes and timelines for making results available and ensuring student anonymity;
- Decisions and communication about courses for which results will not be made available (due to small enrolments or other considerations), and the possible provision of a disclaimer for results that are based on a low response rate;
- Decisions about where to post results and how to determine access to results;
- Decision about the period of time during which results will be available.

When it comes to the actual implementation of the questionnaire by the Faculties at York University, the Joint Subcommittee would like to make the following additional observations:

- That those responsible for the implementation of the questionnaire understand that answers with the number "9" (or "not applicable") should not be computed or disseminated. Instead, they should be treated as missing values. The number "9" (or "not

applicable”) was included in the questionnaire as a possible answer to account for differences across Faculties and course types. For instance, when compared to lectures from the Faculty of Liberal Arts and Professional Studies, some studio courses in Fine Arts were found to employ a totally different course organization, and make use of different course activities and materials. The answer “not applicable” (or number “9” in the scale) was included in the questionnaire exactly to account for these differences. However, its results should not be included in the aggregated average assigned to different courses (or disseminated), as they are not a measure of course evaluation; on the contrary, they stem from the need to create an instrument that works across faculties and course types.

Markus Biehl, Simone Bohn, J. Paul Grayson, Tamara Kelly, Susan Murtha, Alice Pitt, Duff Waring, Diane Woody

APPENDIX A – LIST OF SOURCES CONSULTED AND DISCUSSED

Clayson, D. and D. A. Haley, 2011. “Are students telling us the truth? A critical look at the student evaluation of teaching.” *Marketing Education Review*, vol. 21, no. 2 (Summer), pp. 101-112.

Gravestock, P. and E. Gregor-Greenleaf, 2008. *Student Course Evaluations: Research, Models and Trends*. Toronto: Higher Education Quality Council of Ontario. <http://www.heqco.ca/en-CA/Research/Research%20Publications/Pages/Summary.aspx?link=41&title=Student%20Course%20Evaluations:%20Research,%20Models%20and%20Trends>

Grayson, J. P. , 2009. “Language Background, Ethno-Racial Origin, and Academic Achievement of Students at a Canadian University.” *International Migration*. Vol. 47 (2).

Griffin, B. W. , 2004. “Grading leniency, grade discrepancy, and student ratings of instruction.” *Contemporary Educational Psychology* 29, pp. 410-425. Available online at www.sciencedirect.com

Grimes, P. W., M. J. Meghan and T. W. Woodruff, 2004. “Grades—who’s to blame? Student Evaluation of teaching and locus of control.” *Journal of Economic Education* (Spring), pp. 129-147.

Kember, D. and P. Ginns, 2012. *Evaluating Teaching and Learning: A Practical Handbook for Colleges, Universities and the Scholarship of Teaching*. London & New York: Routledge. (see also item bank at <http://www.routledge.com/books/details/9780415598859/>)

Nowell, C., L. R. Gale and B. Handley, 2010. “Assessing faculty performance using student evaluations of teaching in an uncontrolled setting.” *Assessment and Evaluation in Higher Education*, vol. 35, no. 4 (July), pp. 463-475.

SEEQ (Student Evaluation of Educational Quality), developed by Herbert Marsh, used at many institutions, e.g. Mount Allison and the University of Manitoba.

http://www.mta.ca/pctc/TONI_SEEQ/what_is_seeq.htm

http://intranet.umanitoba.ca/academic_support/catl/media/seeq_booklet.pdf

Vaillancourt, T., 2012. “Students aggress against professors in reaction to receiving poor grades: An effort moderated by student narcissism and self-esteem.” *Aggressive Behavior*, vol. 00, pp. 1-14. Published online in *Wiley Online Library* (wileyonlinelibrary.com).

APPENDIX B – LIST OF COURSE EVALUATIONS FROM YORK UNIVERSITY THAT WERE REVIEWED

Faculty of Education

Faculty of Environmental Studies

Faculty of Fine Arts

Faculty of Health

Faculty of Liberal Arts and Professional Studies: standard version used in most departments, with special forms used in the School of Administrative Studies, Equity Studies, Human Resource Management, and Social Work.

Faculty of Science and Engineering

Glendon College: Departments of French Studies, English, History, International Studies, Mathematics, Multidisciplinary Studies, Political Science and the School of Translation.

Schulich School of Business

Note: No questionnaires from the following three Faculties were reviewed: Lassonde (which had not yet developed its own questionnaires); Osgoode (which falls outside the scope of this joint Employer YUFA project); and the Faculty of Graduate Studies (where practices vary widely and class size may preclude evaluations).

APPENDIX C – FOCUS GROUP STUDY PLAN AND REPORT

CONTENTS OF APPENDIX C

C.1: PURPOSE

C.2: INVESTIGATORS

C.3: RESEARCH METHOD AND PROTOCOL

C.4: ETHICS

C.5: SAMPLING FRAME

C.6: INITIAL QUESTIONNAIRE

C.7: DYNAMICS AND OUTCOMES OF THE GROUPS

C.8: POST FOCUS GROUPS QUESTIONNAIRE

C.1 PURPOSE OF FOCUS GROUPS

The purpose of the questionnaire is to measure how well a course (not the instructor) has worked from the point of view of the students in the course. The set of questions need to work:

- across different course formats;
- across different levels of studies (undergraduate versus graduate);
- across different faculties; and
- across different modes of administration (online and paper).

The purpose of the focus groups is to ensure that

- the resulting questions make sense for evaluating a course;
- the students share a common understanding of what the course-based questions mean to them; and
- the questions work across different types of course formats, levels of study and modes of administration.

It is expected that the questions will be refined during the conduct of the focus groups. The desired output of the focus group exercise is a set of questions that fulfills the above requirements.

C.2 INVESTIGATORS

Principal Investigators (PIs):

- Professors Simone Bohn and Markus Biehl

Institute for Social Research, YorkU:

- Responsible for conducting the focus groups while further developing the formulation of the questions and their response options

C.3 RESEARCH METHOD AND PROTOCOL

Method

Focus groups of 8-10 people each

Sample Frame and Selection

Section C.5 below lists the detailed sampling frame, showing the number of course students (not unique students) enrolled in the different course formats and Faculties (first two tables in C.5). Students are to be sampled with a likeliness that corresponds to the relative frequency with which a particular course type is run. At a minimum, two students are needed per course format across the focus groups run at a particular course level (graduate versus undergraduate). As well, to ensure that the questions can be understood by ESL speakers, at least two ESL speakers need to be contained in the sample (preferably in the undergraduate sample, lecture format).

Separate focus groups are needed for undergraduate and graduate students. For undergraduates, sampling is to occur only for students registered in 2nd year courses or higher, given that new students are likely not yet acclimatized. There is a need to ensure that confirmed student participants are not new to York (i.e. with less than a full semester completed). For graduate students, masters and PhD students can be mixed.

For each block of course formats (see C.5 below) 2 focus groups will be run for undergraduate and one for graduate students for a total of up to 60 students across 6 groups. The two graduate groups will be run after the first set of undergraduate groups.

Questionnaire

The survey questions to be tested are shown below. It is left to ISR to define appropriate response scales or categories (the default being a 7-point Likert scale). The response categories, as well as any further improvements on question wording are to be cleared with the PIs.

Protocol

ISR will employ a professional facilitator to conduct the focus groups. After obtaining informed consent, breaking the ice and providing a context, the facilitator will show students a mock-up version of the survey (on a screen) and ask them to respond to the questions while thinking of a course they have recently taken or are currently taking. The facilitator will ensure that students are considering the course format they were recruited to represent

The facilitator will then ask overview questions and methodically start to check on each of the questions with regards to

- what the students understood the question to mean,
- how relevant they think the question is for evaluating a course, and
- what an improved formulation could look like if the meaning was unclear or inconsistent among different students.

The facilitator will also check whether the questions appear necessary within each category, and whether each category appears necessary from a student point of view.

For the contextual questions about gender and language, the facilitator will check the following:

- Are students comfortable with being asked each of these questions?
- What do students think about the inclusion of these questions in terms of their relevance for the evaluation of the course?

The focus groups will conclude with closing questions, including whether anything major was missed or whether any of the questions appears unimportant or irrelevant.

Either Prof. Bohn or Prof. Biehl will observe each focus group to obtain a sense of potential difficulties with the questionnaire. The questionnaire will be improved after each focus group and then re-tested with the next group. The goal is to have, after 6 rounds of focus groups, a set of questions and response categories that is very likely to work for the range of students, Faculties and course formats relevant to York University.

C. 4 ETHICS

The HPRC Committee has informed the study coordinators that a formal Ethics approval is not required. However, informed consent will be obtained from focus group participants.

C.5 SAMPLING FRAME

NUMBER OF MEETS STARTING IN F13, RELEVANT FORMATS ONLY, BY FOCUS GROUP BLOCK

Start F13
Level (Multiple) UNDERGRADUATE (Levels 2-4)

Count of Coun!	Faculty	AP	ED	ES	FA	GL	HH	LE	LW	SB	SC	Grand Total
1	PRAC	4	52		44	1	18	3			18	140
	STDO				765							765
	WKSP			1	1							2
2	INTR	49	4		2	5	9					69
	LECI	7			3						1	11
	LECT	893	245	12	75	427	247	111	67	148	286	2,511
	LGCL	122				9						131
	SEMR	562		31	22	23	48	1	28		4	719
Grand Total		1,637	301	44	912	465	322	115	95	148	309	4,348

NUMBER OF MEETS STARTING IN F13, RELEVANT FORMATS ONLY, BY FOCUS GROUP BLOCK

Start F13
Level (Multiple) GRADUATE (Levels 5-9)

Count of Coun!	Faculty	ES	GS	LW	SB	Grand Total
1	CLIN		3	12		15
	PRAC		31			31
	STDO		24			24
	WKSP	3	3			6
2	INTR		36			36
	LECI		1			1
	LECT	3	185	3	215	406
	SEMR	37	505	28	12	582
Grand Total		43	788	43	227	1,101

PROPORTIONAL NUMBER OF MEETS STARTING IN F13, RELEVANT FORMATS ONLY, BY FOCUS GROUP BLOCK

Start F13
Level (Multiple I UNDERGRADUATE (Levels 2-4))

Sum of Count	Faculty	ED	ES	FA	GL	HH	LE	LW	SB	SC	Grand Total
FocusBlock	Format	AP									
1	PRAC		1%		1%						2%
	STDO				18%						18%
	WKSP										0%
2	INTR	1%									1%
	LECT										0%
	LECT	21%	6%		2%	10%	6%	3%	2%	3%	57%
	LGCL	3%									3%
	SEMR	13%		1%	1%	1%	1%	1%			16%
Grand Total		37%	7%	1%	21%	10%	7%	3%	2%	3%	98%

PROPORTIONAL NUMBER OF MEETS STARTING IN F13, RELEVANT FORMATS ONLY, BY FOCUS GROUP BLOCK

Start F13
Level (Multiple I GRADUATE (Levels 5-9))

Sum of Count	Faculty	GS	LW	SB	Grand Total	
FocusBlock	Format	ES				
1	CLIN		1%		1%	
	PRAC		3%		3%	
	STDO		2%		2%	
	WKSP				0%	
2	INTR		3%		3%	
	LECT				0%	
	LECT		17%	20%	36%	
	SEMR	3%	46%	3%	53%	
Grand Total		3%	71%	4%	21%	99%

NOTE: At least two students are needed for each course type.

C.6 INITIAL QUESTIONNAIRE

Instructions for Focus Groups Mediators

- Please ask whether any significant information is missing for evaluating a COURSE.
- Please ask which question(s) students are most likely to draw upon when selecting a course.

Instructions to Respondents

The results of the course evaluation questions will be made public. Answers to these questions will help the Faculty further improve the course and curriculum.

Contextual Questions

- Was this course mandatory for you? (Y/N)
- Compared to other courses, how much effort did you expend for this course? (less ... more)
- FOR TESTING ONLY: I am ... (male/female/other)
- FOR TESTING ONLY: My first language is English (Y/N)

Course Evaluation Questions

Organization

1. A roadmap of the course (i.e. a syllabus, course outline, a breakdown with topics, assignments and exams etc.) was provided.
2. The roadmap presented was followed.
3. The course's learning outcomes (what you will be able to do, know, value at the end of the course) were communicated.
Possible alternative to be tested: The course's objectives (what you will be able to do, know, value at the end of the course) were communicated.
4. Expectations and requirements for assignments were communicated.
5. Policies and procedures such as attendance, participation, missed tests, late assignments, contacting instructors etc. were stated in the course outline.

Course Materials

6. The course materials (readings, course kits, textbooks, audio-visual materials, lab manuals, websites etc.) were directly related to the course's learning outcomes.
Possible alternative to be tested: The course materials (readings, course kits, textbooks, audio-visual materials, lab manuals, websites etc.) were directly related to the course's objectives.
7. The course materials and resources helped me achieve the course's learning outcomes.
Possible alternative to be tested: The course materials and resources helped me achieve the course's objectives.

Learning activities

8. The course activities (discussions, simulations, exercises, presentations etc.) helped me achieve the learning outcomes (alternatively: course objectives).
9. The course assignments and tests were directly related to the course's learning outcomes.
Possible alternative to be tested: The course assignments were directly related to the course's objectives.

Learning Experience

10. The course helped me grow intellectually.

C.7 FOCUS GROUPS: DYNAMICS AND OUTCOMES

Following the plan spelled out in the Study Plan, the Institute of Social Research (ISR) at York University recruited graduate and undergraduate students to participate in six focus groups. The selection of students was based primarily on the format of courses that they had taken recently, and on their level of study (graduate vs. undergraduate studies). Care was also taken to assure that students from different Faculties were selected for the study, and that both male and female students were recruited.

The primary goals of these focus groups were, first, to check whether students deemed the questions devised by the Joint Subcommittee on Course Evaluations appropriate for the purpose of evaluating courses at York University. The second objective was to gauge whether students' understanding of the questions proposed corresponded to the overall meaning attributed to them by the Joint Subcommittee.

The focus groups took place on the following days: November 5, 6, 7, 12, 13 and 14, 2013. Each session lasted approximately 90 minutes. Professor Markus Biehl and Professor Simone Bohn took turns observing each of the focus groups. The latter were moderated by John Pollard, from the ISR. Overall, 42 students participated in this study.

Several issues emerged from the focus group dynamics. In each instance, the questionnaire referenced is the initial questionnaire (see section C.6 above).

a.) *The need to replace the term “learning outcomes” with “course objectives”.*

Graduate students easily recognized the term “learning outcomes”, and were able to describe it in ways that match the meaning attributed to the notion by the Joint Subcommittee. However, undergraduate students overwhelmingly reported either not being familiar with the term or not understanding its meaning. Thus, the Joint Subcommittee decided to replace “learning outcomes” with “course objectives”. The final three focus groups worked with the latter term, instead of the former, and there was a substantial improvement in understanding from the part of undergraduate students.

b.) *The need to harmonize the seven-point scale used to answer the course evaluation questions.*

The students who participated in the study considered it important to have the option “not sure/not applicable” for every single question. The Joint Subcommittee agreed to take this suggestion, as it became evident that there is a great diversity of both course formats at York University and teaching and learning cultures across Faculties.

c.) *The need to separate laboratories and tutorials from the remaining course activities.*

Students considered labs and tutorials to have a structure different from of other course activities listed on the questionnaire, such as “lectures, discussions, assignments, exercises and presentations, etc.” They claimed that a strongly positive evaluation of one type of course activity did not necessarily imply a similar view regarding the other. Given this assessment, the Joint Subcommittee decided to have two distinct sets of questions for these course activities.

d.) *The need to include assignments that are specific to fourth-year undergraduate courses and graduate courses.*

One of the questions asked students to evaluate the degree to which “the course tests and exams were directly related to the course objectives.” Upper-level undergraduate students and graduate students remarked that their courses hardly ever use tests and exams as the final course assignment. This consideration prompted the Joint Subcommittee to change the formulation above to “the course tests/exams or final paper/essay were directly related to the course objectives”.

e.) *The need to expand the dichotomy “mandatory vs. elective” courses to better capture the dynamics of some programs across the university.*

One of the background questions that students were invited to contemplate was whether the course under evaluation was mandatory. In this regard, in the course of the focus groups, the Joint Subcommittee learned that, in the case of some programs, this dichotomy is not as clear cut. Some academic units have a fixed set of mandatory courses, a list of required courses from which students can choose, and some additional credits that allow students to select whatever courses they desire to pursue. In order to ensure that the questionnaire under study encompasses the experience of these particular programs, the first background question was changed as follows: “Was this course mandatory for you? () 1. Yes; () 2. Yes, as a choice from a list of required courses; or () 3. No.”

f.) *The need to eliminate the background questions regarding gender and language.*

Participants in the focus groups deemed these questions to be highly inappropriate and to be designed with the intent of discrediting students’ negative feedback. Some argued that the gender question might be offensive to sexual minorities, and that students (i.e. future questionnaire respondents) might take the language question as a proxy for gauging ethnicity and culture. Given

these assessments, the Joint Subcommittee decided against maintaining these two questions in the questionnaire.

The post-focus-groups questionnaire is presented in section C.8 below. See Appendix D for a discussion of the pre-test that was conducted subsequently, in both web-based and paper-based formats.

C.8 QUESTIONNAIRE – AFTER FOCUS GROUPS

For each of the questions below, please choose an answer based on this scale:						
Strongly agree	Agree	Agree somewhat	Disagree somewhat	Disagree	Strongly disagree	Not sure/ Not applicable
1	2	3	4	5	6	7
Question						Your answer
1) A comprehensive roadmap of the course (i.e., a syllabus or course outline, a breakdown with topics, assignments and exams etc.) was provided.						
2) The roadmap provided was mostly followed.						
3) Expectations and requirements for assignments were clear.						
4) Important policies and procedures (e.g., attendance, participation, missed tests, late assignments, contacting instructors, etc.) were stated in the course outline.						
5) The course objectives (i.e., what you are expected to know and be able to do at the end of the course) were communicated.						
6) The course materials (e.g., course kits, textbooks, readings, audio visual materials, lab manuals, websites, etc.) were directly related to the course objectives.						
7) The course materials helped me achieve the course objectives.						
8) The course activities (e.g. lectures, discussions, simulations, assignments, exercises and presentations, etc.) helped me achieve the course objectives.						
9) The labs or tutorials helped me achieve the course objectives.						
10) The course tests/exams or final paper/essay were directly related to the course objectives.						
11) The course helped me grow intellectually.						

Background questions

- a) Was this course mandatory for you?
 1. Yes; 2. Yes, as a choice from a list of required courses, or 3. No
- b) Approximately how many lectures/seminars/sessions of this course did you attend (including online “attendance” in the case of online courses)?
 1. All of them; 2. Most of them; 3. More than half of them; 4. About half of them; 5. Fewer than half of them.

APPENDIX D – PSYCHOMETRIC TESTING.

CONTENTS OF APPENDIX D

D.1: PURPOSE OF PRE-TEST

D.2: METHOD, SAMPLING FRAME, GUIDELINES AND PROTOCOL

D.3: ANALYSIS

D.4: RECOMMENDED INSTRUCTION TO STUDENTS AND QUESTIONNAIRE

D.5: SUMMARY OF STATISTICAL ANALYSIS

D.1 PURPOSE

After carrying out six focus groups and fine-tuning the questionnaire instrument through this method, the Joint Subcommittee on Course Evaluation devoted itself to pre-testing the questionnaire.

The primary objectives of the pre-testing were essentially twofold: to determine the overall reliability and validity of the questionnaire, and to identify and eliminate redundant variables, i.e. questionnaire items that are highly correlated with others and whose presence does not improve the overall explanatory power of the instrument.

D.2 METHOD, SAMPLING FRAME, GUIDELINES AND PROTOCOL

Following the same underlying logic of the organization of the focus groups, the Joint Subcommittee selected different types of courses from distinct Faculties. Given that the final questionnaire will work as a pan-university instrument, the reasoning behind this selection was to ensure that the final course evaluation questions are applicable to the diverse course formats found in different Faculties. In addition, the Subcommittee decided to carry out the pre-testing in both an electronic and a paper-based modes of delivery. The table below summarizes the number of students and courses across Faculties that were included in the pretesting.

Table 1. Courses and number of students included in the pretesting, by mode of delivery

<i>Paper-based evaluations</i>		
Faculty	Number of courses	Total number of students
Fine Arts	1	112
Liberal Arts and Professional Studies	2	119
Science	1	148
<i>Online evaluations</i>		
Faculty	Number of courses	Total number of students
Liberal Arts and Professional Studies	1	8
Health	1	288
Schulich	2	101
Science	1	105
Total	Number of courses	Total number of students

All Faculties	9	881
----------------------	---	-----

After the paper-based course evaluations were administered in four different classrooms, the Institute of Social Research entered the data into a spreadsheet, which was later merged with the data obtained through the online method. Three members of the Joint Subcommittee carried out independent statistical analysis of the database. Care was also taken to verify how well the instrument worked when administered in class (on paper) and online.

D.3 ANALYSIS

The main conclusions of the analyses were, first, that the instrument proposed is reliable and valid, and, second, that some questions can be easily trimmed off the final version of the questionnaire, as they correlated highly with other items, without adding substantially to the instrument's overall explanatory power. Third, it was determined that the proposed questionnaire provides reliable results in both an online mode of delivery, and in paper-based mode. Similarly, the instrument also works well across Faculties and course types.

The final proposed questionnaire and instruction to students are provided below in section D.4. Following that, Section D.5 provides the basic statistical analysis that led to the conclusions reached.

D.4 RECOMMENDED INSTRUCTION TO STUDENTS AND QUESTIONNAIRE

Dear Student,

With the objective of aiding upcoming students in their course selection, York University has developed a common set of course evaluation questions that will be used across all its Faculties and course types. To that purpose, the class average of the answers to the seven course evaluation questions below will be made public.

Thank you for your participation in this important work!

For each of the questions below, please choose an answer based on this scale:							
Strongly disagree	Disagree	Somewhat Disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree	Not applicable
1	2	3	4	5	6	7	9
Question							Your answer
1) A comprehensive roadmap of the course (i.e., a syllabus or course outline, a breakdown with topics, assignments and exams etc.) was provided.							
2) The roadmap provided was mostly followed.							
3) Important policies and procedures (e.g., attendance, participation, missed tests, late assignments, contacting instructors, etc.) were stated in the course outline.							
4) The course materials (e.g., course kits, textbooks, readings, audio visual							

materials, lab manuals, websites, etc.) helped me achieve the course objectives.	
5) The course activities (e.g. lectures, discussions, simulations, assignments, exercises and presentations, etc.) helped me achieve the course objectives.	
6) The course tests/exams or final paper/essay were directly related to the course objectives.	
7) The course helped me grow intellectually.	

Background questions

a) Was this course mandatory for you?

- 1. Yes;
- 2. Yes, as a choice from a list of required courses;
- 3. No

b) Approximately how many lectures/seminars/sessions of this course did you attend (including online “attendance” in the case of online courses)?

- 1. All of them
- 2. Almost all of them
- 3. More than half of them
- 4. About half of them
- 5. Fewer than half of them.

D.5 SUMMARY OF THE STATISTICAL ANALYSIS

Questionnaire items (outcome of the focus groups) and their respective variable names in the statistical analysis

Questionnaire item	Variable name
A comprehensive roadmap of the course (i.e., a syllabus or course outline, a breakdown with topics, assignments and exams etc.) was provided.	<i>road_prov</i>
The roadmap provided was mostly followed.	<i>road_foll</i>
Expectations and requirements for assignments were clear.	<i>exp_clear</i>
Important policies and procedures (e.g., attendance, participation, missed tests, late assignments, contacting instructors, etc.) were stated in the course outline.	<i>policies_stated</i>
The course objectives (i.e., what you are expected to know and be able to do at the end of the course) were communicated.	<i>obj_communic</i>
The course materials (e.g., course kits, textbooks, readings, audio visual materials, lab manuals, websites, etc.) were directly related to the course objectives.	<i>mats_related</i>
The course materials helped me achieve the course objectives.	<i>mats_helped</i>
The course activities (e.g. lectures, discussions, simulations, assignments, exercises and presentations, etc.) helped me achieve the course objectives.	<i>act_helped</i>
The labs or tutorials helped me achieve the course objectives.	<i>labs_tut</i>
The course tests/exams or final paper/essay were directly related to the course objectives.	<i>tests_related</i>
The course helped me grow intellectually.	<i>grow_int</i>
Was this course mandatory for you?	<i>mandatory</i>
Approximately how many lectures/seminars/sessions of this course did you attend (including online “attendance” in the case of online courses)?	<i>Attend</i>

OVERVIEW STATISTICS

Description of Data

variable	mean	sd	min	p25	p50	p75	max	skewness
road_prov	5.37	1.00	0	5	6	6	6	-2.40
road_foll	5.16	1.14	0	5	5	6	6	-2.27
exp_clear	5.00	1.20	0	4	5	6	6	-1.46
policies_s~d	5.43	0.98	0	5	6	6	6	-2.79
obj_communic	5.00	1.08	0	5	5	6	6	-1.49
mats_related	5.18	1.10	0	5	5	6	6	-2.07
mats_helped	4.87	1.27	0	4	5	6	6	-1.63
act_helped	4.87	1.32	0	4	5	6	6	-1.54
tests_rela~d	4.77	1.61	0	4	5	6	6	-1.91
grow_int	4.78	1.33	0	4	5	6	6	-1.48
mandatory	0.71	0.37	0	.5	1	1	1	-0.88
labs_tut	4.32	1.74	0	4	5	6	6	-1.25
attend	0.88	0.15	.25	.8	1	1	1	-1.59

```
labs_tut_c~r | 4.06 1.97 0 4 5 6 6 -1.05
```

Given the response items, it could be expected that the items are negatively skewed.

Differences in select representative variables by Course Format, Faculty, Survey Method

Format	stats	road_p~v	mats_h~d	grow_int
LECI	mean	5.75	5.5	5.25
	sd	.707	.756	.707
	N	8	8	8
LECT	mean	5.47	4.61	4.56
	sd	.691	1.04	1.35
	N	72	72	72
LECT+LAB	mean	5.61	5.26	5.04
	sd	.648	.951	1.08
	N	497	497	496
LECT+TUTR	mean	4.88	4.19	4.3
	sd	1.36	1.54	1.57
	N	276	276	276
SEMR	mean	5.46	5.25	5.14
	sd	1.26	1	1.48
	N	28	28	28
Total	mean	5.37	4.87	4.78
	sd	1.01	1.27	1.33
	N	881	881	880

Faculty	stats	road_p~v	mats_h~d	grow_int
AP	mean	5.46	4.5	4.41
	sd	.966	1.46	1.62
	N	127	127	127
FA	mean	4.31	4.13	4.16
	sd	1.57	1.5	1.6
	N	112	112	112
HH	mean	5.76	5.4	5.21
	sd	.488	.87	.911
	N	288	288	288
SB	mean	5.16	4.29	4.29
	sd	1.07	1.41	1.43
	N	101	101	101
SC	mean	5.43	5.02	4.93
	sd	.771	1.07	1.2
	N	253	253	252
Total	mean	5.37	4.87	4.78
	sd	1.01	1.27	1.33
	N	881	881	880

SurveyType	stats	road_p~v	mats_h~d	grow_int
Online	mean	5.54	5.06	4.99
	sd	.764	1.14	1.12
	N	502	502	502
Paper	mean	5.14	4.63	4.49
	sd	1.23	1.39	1.53
	N	379	379	378
Total	mean	5.37	4.87	4.78
	sd	1.01	1.27	1.33
	N	881	881	880

CORRELATION ANALYSIS

Helps identify candidate variables that can potentially be deleted as they measure essentially the same thing; high correlations marked bold.

	road_p~v	road_f~l	exp_cl~r	polici~d	obj_co~c	mats_r~d	mats_h~d
road_prov	1.0000						
road_foll	0.6548*	1.0000					
exp_clear	0.4873*	0.4248*	1.0000				
policies_s~d	0.4677*	0.4233*	0.4317*	1.0000			
obj_communic	0.4374*	0.4135*	0.5121*	0.4446*	1.0000		
mats_related	0.4736*	0.4497*	0.4997*	0.3871*	0.5152*	1.0000	
act_helped	0.4189*	0.3715*	0.4968*	0.3442*	0.5656*	0.5515*	0.7055*
tests_rela~d	0.2737*	0.2293*	0.3099*	0.2166*	0.3852*	0.3674*	0.3951*
grow_int	0.4054*	0.3889*	0.4327*	0.3406*	0.4879*	0.5089*	0.5488*
mandatory	-0.0795	-0.0425	-0.0796	-0.0533	-0.0021	-0.0700	-0.0873
attend	0.0334	0.0882*	0.0495	0.0364	0.0474	-0.0122	0.0439
labs_tut_c~r	0.1153*	0.1265*	0.2822*	0.1607*	0.2299*	0.2478*	0.2932*

	act_he~d	tests_~d	grow_int	mandat~y	attend	labs_t~r
act_helped	1.0000					
tests_rela~d	0.3785*	1.0000				
grow_int	0.5895*	0.4383*	1.0000			
mandatory	-0.0931*	0.0739	-0.0470	1.0000		
attend	0.0876	0.0936*	0.0954*	0.1568*	1.0000	
labs_tut_c~r	0.3551*	0.2457*	0.2451*	0.0852	0.0015	1.0000

Candidates: road_prov - road followed; materials helped - materials helped; grow intellectually - materials related, materials helped, activities helped; activities helped - objectives communicated, materials related, materials helped;

Overall Reliability

Chronbach's Alpha - all items

To establish internal consistency (>= 0.7 = good)

Test scale = mean(unstandardized items)

Item	Obs	Sign	item-test correlation	item-rest correlation	average interitem covariance	alpha
road_prov	881	+	0.6509	0.5823	.4539552	0.8495
road_foll	881	+	0.6210	0.5382	.4516307	0.8516
exp_clear	879	+	0.7030	0.6295	.434133	0.8457
policies_s~d	878	+	0.5624	0.4842	.467633	0.8544
obj_communic	879	+	0.7286	0.6677	.4384998	0.8445
mats_related	877	+	0.7321	0.6695	.4347077	0.8436
mats_helped	881	+	0.7744	0.7106	.4158049	0.8397
act_helped	877	+	0.7876	0.7242	.4108061	0.8388
tests_rela~d	869	+	0.5879	0.4566	.4351296	0.8598
grow_int	880	+	0.7286	0.6499	.4207099	0.8438
mandatory	862	-	0.0965	0.0573	.5288497	0.8684
labs_tut	877	+	0.5525	0.4040	.4410673	0.8673
attend	865	+	0.0887	0.0733	.5305596	0.8678
Test scale					.4510675	0.8622

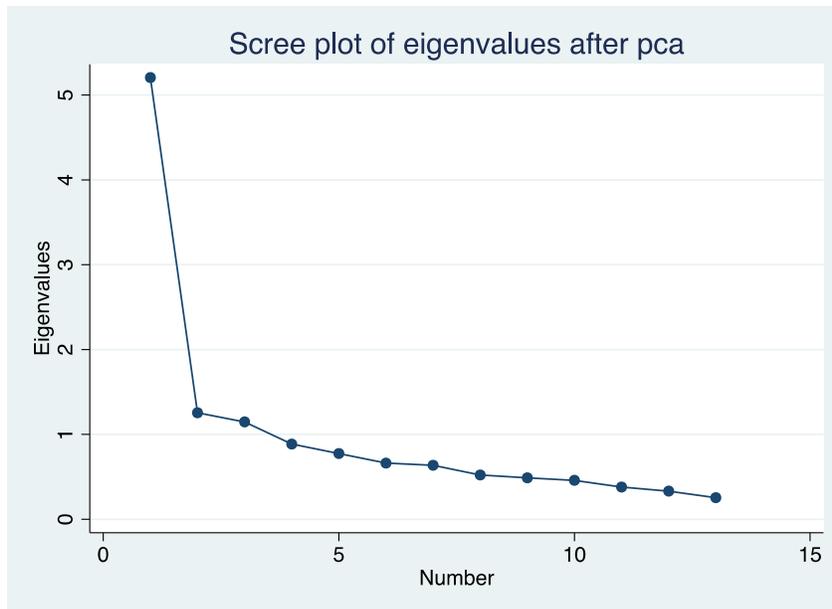
EXPLORATORY FACTOR ANALYSIS

Determines linearly uncorrelated factors (components), each containing a set of variables that are structurally similar; components reveal the data's internal structure.

First Iteration

Principal components/correlation	Number of obs	=	835
	Number of comp.	=	3
	Trace	=	13
Rotation: (unrotated = principal)	Rho	=	0.5853

Component	Eigenvalue	Difference	Proportion	Cumulative
Comp1	5.2061	3.95135	0.4005	0.4005
Comp2	1.25476	.107079	0.0965	0.4970
Comp3	1.14768	.261836	0.0883	0.5853
Comp4	.885842	.111186	0.0681	0.6534
Comp5	.774656	.113229	0.0596	0.7130
Comp6	.661427	.0252913	0.0509	0.7639
Comp7	.636136	.114597	0.0489	0.8128
Comp8	.521539	.0334479	0.0401	0.8529
Comp9	.488091	.0291585	0.0375	0.8905
Comp10	.458933	.079445	0.0353	0.9258
Comp11	.379488	.048067	0.0292	0.9550
Comp12	.331421	.0774929	0.0255	0.9805
Comp13	.253928	.	0.0195	1.0000



```
Principal components/correlation          Number of obs   =      835
                                          Number of comp. =       3
                                          Trace           =      13
Rotation: orthogonal entropy (Kaiser off) Rho           =     0.5853
```

Component	Variance	Difference	Proportion	Cumulative
Comp1	4.01653	1.63203	0.3090	0.3090
Comp2	2.3845	1.17699	0.1834	0.4924
Comp3	1.20751	.	0.0929	0.5853

Rotated components (blanks are abs(loading)<.2)

Variable	Comp1	Comp2	Comp3	Unexplained
road_prov		0.5331		.296
road_foll		0.5452		.3215
exp_clear	0.2278	0.2315		.466
policies_s~d		0.4201		.507
obj_communic	0.2795			.4285
mats_related	0.3247			.387
mats_helped	0.4009			.3125
act_helped	0.4172			.3068
tests_rela~d	0.3579			.5499
grow_int	0.3518			.4375
mandatory			0.6905	.4025
attend			0.6713	.432
labs_tut_c~r	0.4159	-0.3128		.5444

Three components: (1) Course materials and activities; (2) Course structure; (3) Contingency variables

Cross-loading items do not fit into one or the other component -> candidates for deletion, if theoretically justifiable.

Second Iteration

```
Principal components/correlation      Number of obs   =      839
                                      Number of comp. =       3
                                      Trace             =      11
Rotation: orthogonal entropy (Kaiser off)  Rho             =    0.6263
```

Component	Variance	Difference	Proportion	Cumulative
Comp1	3.50584	1.32402	0.3187	0.3187
Comp2	2.18182	.980034	0.1983	0.5171
Comp3	1.20179	.	0.1093	0.6263

Rotated components (blanks are abs(loading)<.2)

Variable	Comp1	Comp2	Comp3	Unexplained
road_prov		0.5676		.2778
road_foll		0.5984		.2749
policies_s~d		0.4942		.4568
obj_communic	0.3153			.4216
mats_related	0.3664			.3688
mats_helped	0.4580			.2715
act_helped	0.4532			.2947
tests_rela~d	0.4218		0.2033	.4909
grow_int	0.4148			.3933
mandatory			0.6964	.4091
attend			0.6680	.4511

- Possible to delete onj_communic? Highly correlated w/ materials ... activities
- Possible to delete mats_related? Very highly correlated w/ mats_helped.

Third Iteration

```
Principal components/correlation      Number of obs   =      844
                                      Number of comp. =       3
                                      Trace           =       9
Rotation: orthogonal entropy (Kaiser off)  Rho             =    0.6459
```

Component	Variance	Difference	Proportion	Cumulative
Comp1	2.59721	.573541	0.2886	0.2886
Comp2	2.02366	.831594	0.2249	0.5134
Comp3	1.19207	.	0.1325	0.6459

Rotated components (blanks are abs(loading)<.2)

Variable	Comp1	Comp2	Comp3	Unexplained
----------	-------	-------	-------	-------------

road_prov			0.5821		.262
road_foll			0.6106		.2662
policies_s~d			0.5089		.4588
mats_helped		0.5025			.2871
act_helped		0.5166			.2695
tests_rela~d		0.4902			.4552
grow_int		0.4863			.3463
mandatory			0.7139		.384
attend			0.6656		.458

Clean set of components; further deletions not recommended since conceptually the questions are all very different from one another or logically follow upon each other; combining questions not possible. Note: Component 3 not a factor in itself, but merely the set of contingency questions.

Test of Survey with 7+2 Questions

Chronbach's ALPHA

Test scale = mean(unstandardized items)

Item	Obs	Sign	item-test correlation	item-rest correlation	average interitem covariance	alpha
road_prov	881	+	0.6827	0.5811	.3650187	0.7627
road_foll	881	+	0.6589	0.5363	.3627516	0.7679
policies_s~d	878	+	0.5698	0.4479	.3912624	0.7787
mats_helped	881	+	0.7634	0.6547	.3228451	0.7469
act_helped	877	+	0.7763	0.6673	.3176578	0.7460
tests_rela~d	869	+	0.6173	0.4127	.3500127	0.7942
grow_int	880	+	0.7523	0.6317	.3219833	0.7505
mandatory	862	-	0.1145	0.0553	.4785133	0.8091
attend	865	+	0.1141	0.0909	.4803469	0.8075
Test scale					.3767924	0.7961

All items OK, overall construct reliable.

Confirmatory Factor Analysis - Single Factor

Structural Equations Model

```
. sem (STR -> road_prov road_foll policies_stated mats_helped act_helped tests_related
grow_int), latent( STR) stand method(mlmv)
```

```
Structural equation model          Number of obs      =      881
Estimation method = mlmv
Log likelihood      = -8995.4606
```

```
( 1) [road_prov]STR = 1
```

Standardized	Coef.	OIM Std. Err.	z	P> z	[95% Conf. Interval]
Measurement					
road_prov <-					

STR		.6167629	.0253255	24.35	0.000	.5671259	.6664
_cons		5.320481	.1311512	40.57	0.000	5.063429	5.577532

road_foll <-							
STR		.5795153	.0267122	21.69	0.000	.5271603	.6318703
_cons		4.547648	.1134564	40.08	0.000	4.325277	4.770018

policies_stated <-							
STR		.4997593	.0287465	17.39	0.000	.4434171	.5561015
_cons		5.540994	.1363882	40.63	0.000	5.273678	5.80831

mats_helped <-							
STR		.7839417	.017532	44.71	0.000	.7495795	.8183038
_cons		3.830447	.0972736	39.38	0.000	3.639794	4.0211

act_helped <-							
STR		.7961754	.017182	46.34	0.000	.7624994	.8298515
_cons		3.689257	.0943408	39.11	0.000	3.504353	3.874162

tests_related <-							
STR		.4965917	.0285155	17.41	0.000	.4407023	.552481
_cons		2.946168	.0784605	37.55	0.000	2.792388	3.099948

grow_int <-							
STR		.7202957	.0197836	36.41	0.000	.6815206	.7590708
_cons		3.573483	.0918821	38.89	0.000	3.393397	3.753568

Variance							
e.road_prov		.6196035	.0312396			.561303	.6839594
e.road_foll		.664162	.0309603			.6061705	.7277015
e.policies_stated		.7502406	.0287327			.6959873	.8087231
e.mats_helped		.3854355	.0274882			.3351556	.4432584
e.act_helped		.3661047	.0273598			.3162228	.423855
e.tests_related		.7533967	.0283211			.6998839	.8110011
e.grow_int		.4811741	.0285001			.4284354	.5404047
STR		1	.			.	.

LR test of model vs. saturated: chi2(14) = 395.79, Prob > chi2 = 0.0000

All loadings OK.

Model Fit Statistics

Fit statistic	Value	Description

Likelihood ratio		
chi2_ms(14)	395.791	model vs. saturated
p > chi2	0.000	
chi2_bs(21)	2300.607	baseline vs. saturated
p > chi2	0.000	

Population error		
RMSEA	0.176	Root mean squared error of approximation
90% CI, lower bound	0.161	
upper bound	0.191	
pclose	0.000	Probability RMSEA <= 0.05

Information criteria		
AIC	18032.921	Akaike's information criterion
BIC	18133.323	Bayesian information criterion

-----+-----			
Baseline comparison			
CFI		0.833	Comparative fit index
TLI		0.749	Tucker-Lewis index
-----+-----			
Size of residuals			
CD		0.861	Coefficient of determination
-----+-----			

Not a well fitting model (CFI, TLI, CD < 0.9, RMSEA > 0.08).

Confirmatory Factor Analysis - Two Factors

Structural Equations Model

```
. sem (STR -> road_prov road_foll policies_stated) (CRS -> mats_helped act_helped
tests_related grow_int), latent( STR CRS) stand method(mlmv)
```

```
Structural equation model                Number of obs      =          881
Estimation method = mlmv
Log likelihood      = -8824.8419
```

- (1) [road_prov]STR = 1
- (2) [mats_helped]CRS = 1

		Standardized	Coef.	OIM Std. Err.	z	P> z	[95% Conf. Interval]
-----+-----							
Measurement							
road_prov <-							
	STR		.8366876	.0188296	44.43	0.000	.7997822 .8735929
	_cons		5.320481	.1311512	40.57	0.000	5.063429 5.577532
-----+-----							
road_foll <-							
	STR		.7728363	.0200036	38.63	0.000	.7336299 .8120427
	_cons		4.547648	.1134564	40.08	0.000	4.325277 4.770018
-----+-----							
policies_stated <-							
	STR		.5678933	.0269698	21.06	0.000	.5150335 .6207531
	_cons		5.539713	.1364006	40.61	0.000	5.272373 5.807054
-----+-----							
mats_helped <-							
	CRS		.8168684	.0157994	51.70	0.000	.7859021 .8478346
	_cons		3.830447	.0972736	39.38	0.000	3.639794 4.0211
-----+-----							
act_helped <-							
	CRS		.8373938	.0151244	55.37	0.000	.8077506 .867037
	_cons		3.69227	.0943337	39.14	0.000	3.507379 3.87716
-----+-----							
tests_related <-							
	CRS		.4968052	.0286198	17.36	0.000	.4407114 .552899
	_cons		2.946815	.0784517	37.56	0.000	2.793052 3.100577
-----+-----							
grow_int <-							
	CRS		.7136217	.0203864	35.00	0.000	.6736651 .7535784
	_cons		3.576834	.091873	38.93	0.000	3.396766 3.756902
-----+-----							
Variance							

e.road_prov		.2999539	.031509			.2441401	.3685275
e.road_foll		.4027241	.0309191			.3464629	.4681214
e.policies_stated		.6774972	.0306319			.6200431	.7402751
e.mats_helped		.3327261	.0258121			.2857937	.3873655
e.act_helped		.2987717	.0253301			.2530311	.3527808
e.tests_related		.7531846	.0284369			.6994615	.811034
e.grow_int		.490744	.0290964			.4369049	.5512176
STR		1	.			.	.
CRS		1	.			.	.

Covariance								
STR								
	CRS		.6343029	.0280372	22.62	0.000	.5793511	.6892548

LR test of model vs. saturated: chi2(13) = 54.55, Prob > chi2 = 0.0000

All loadings OK.

Model Fit

Fit statistic		Value	Description
Likelihood ratio			
chi2_ms(13)		54.554	model vs. saturated
p > chi2		0.000	
chi2_bs(21)		2300.607	baseline vs. saturated
p > chi2		0.000	
Population error			
RMSEA		0.060	Root mean squared error of approximation
90% CI, lower bound		0.044	
upper bound		0.077	
pclose		0.139	Probability RMSEA <= 0.05
Information criteria			
AIC		17693.684	Akaike's information criterion
BIC		17798.867	Bayesian information criterion
Baseline comparison			
CFI		0.982	Comparative fit index
TLI		0.971	Tucker-Lewis index
Size of residuals			
CD		0.961	Coefficient of determination

Two factor model exhibits a good fit => presence of two underlying factors confirmed.

Potential Further Improvements

Modification indices

Could the survey be further improved?

		MI	df	P>MI	EPC	Standard EPC

Measurement						
policies_stated <-	CRS	9.158	1	0.00	.1352518	.1434763
act_helped <-	STR	4.882	1	0.03	-.1457991	-.0933288
grow_int <-	STR	9.405	1	0.00	.2074594	.131282
Covariance						
e.road_prov	e.road_foll	9.158	1	0.00	.1762086	.4428845
e.road_foll	e.act_helped	4.390	1	0.04	-.0540558	-.1041504
e.mats_helped	e.act_helped	32.863	1	0.00	.2942601	.556219
	e.grow_int	18.707	1	0.00	-.1822813	-.2657726
e.act_helped	e.tests_related	10.392	1	0.00	-.1564786	-.1547075
e.tests_related	e.grow_int	22.175	1	0.00	.2420389	.1845846

EPC = expected parameter change

Model fit could be improved if mats_helped and act_helped were allowed to correlate with one another, or if one of the variables was deleted

Fit of model with act_helped deleted:

Fit statistic	Value	Description
Likelihood ratio		
chi2_ms(8)	16.155	model vs. saturated
p > chi2	0.040	
chi2_bs(15)	1567.325	baseline vs. saturated
p > chi2	0.000	
Population error		
RMSEA	0.034	Root mean squared error of approximation
90% CI, lower bound	0.007	
upper bound	0.058	
pclose	0.850	Probability RMSEA <= 0.05
Information criteria		
AIC	15407.728	Akaike's information criterion
BIC	15498.569	Bayesian information criterion
Baseline comparison		
CFI	0.995	Comparative fit index
TLI	0.990	Tucker-Lewis index
Size of residuals		
CD	0.935	Coefficient of determination

Fit better, but conceptually difficult to delete this variable.