52. **EXECUTIVE COMMITTEE ACTION SUMMARIES FOR THE MEETING OF MARCH 29, 1999**

The Executive Committee Action Summaries of March 29, 1999 were before members.

Associate Vice-President and Registrar B Silzer noted, in reference to Minute 50 of the Executive Committee Action Summaries, that the Executive Committee had approved changes to the Academic Schedule that would see, in the future, the University begin classes one day earlier in the fall and that would add a 'floater' day near Remembrance Day. He noted that these changes would help to balance the number of instructional hours between Monday/Wednesday/Friday classes and Tuesday/Thursday classes.

MOTION It was **MOVED** by Vice-President (Academic) and Provost Owram and seconded by Associate Vice-President and Registrar Silzer that the Executive Committee Action Summaries of March 29, 1999 be approved.

CARRIED

53. **REPORT OF THE BOARD OF GOVERNORS**

There was no report from the Board of Governors, as it did not meet in April.

54. **REPORT OF THE GFC NOMINATING COMMITTEE**

There was no written report from the Nominating Committee. Professor F Khanna, Chair, GFC Nominating Committee, noted that a mail ballot regarding the election of an academic staff member to the Academic Planning Committee (APC) had been mailed to members on March 30, 1999. He also noted that additional reports would be sent to members in May.

55. **UNIVERSAL STUDENT RATINGS OF INSTRUCTION: RECOMMENDATION FROM THE GFC TEACHING AND LEARNING COMMITTEE (TLC) FOR REVISIONS TO SECTION 111.3 OF THE GFC POLICY MANUAL (UNIVERSAL STUDENT RATINGS OF INSTRUCTION)**

Members considered a proposal from the GFC Teaching and Learning Committee (TLC) for revisions to Section 111.3 of the GFC Policy Manual (Universal Student Ratings of Instruction), as set out in the right-hand column of the comparative table that was before members.

It was noted that the Executive Committee had considered and endorsed the above-noted proposal at its meeting of March 29, 1999.

Relevant background information and legislation (Section 111.3, GFC Policy Manual) were before members.

Associate Vice-President (Academic) AM Decore, Chair, TLC, and members of TLC, Professor M Price, Professor M Roeder, and Ms Andrusky, were in attendance and presented the proposal.

MOTION It was **MOVED** by Vice-President (Academic) and Provost Owram and seconded by Mr Murphy that GFC approve changes to Section 111.3 of the GFC Policy Manual (Universal Student Ratings of Instruction), as set out in the right-hand column of the comparative table that was before members, with the changes to take effect immediately.

During a lengthy discussion, a number of points were raised, including the following:

- Dr Decore noted that the Association of Academic Staff: University of Alberta (AAS:UA) had indicated they would prefer further consultation before GFC considered the proposal. She indicated reluctance to defer consideration of the proposal for three reasons: 1) AAS:UA had been consulted throughout the process and, during an AAS:UA Executive meeting, had not raised objections of substance; 2) Department chairs had been notified of the proposed changes and many instructors were asked to participate in a recent pilot study; and 3) deferring consideration of the proposed changes until the Summer would have make it difficult for student input at GFC and deferring consideration until the Fall would delay implementation of the changes for another academic year.
- Professors Price and Roeder addressed members' concerns regarding the misuse of statistics, explained the use of the Tukey Box- and-Whisker Plot and summarized a review of the literature which was available on student ratings of instruction.

MOTION TO TABLE It was **MOVED** by Professor Renke and seconded by Professor Kline that GFC **TABLE** the Motion, pending satisfactory consultation with academic staff members. **A motion to TABLE is non-debatable.**

DEFEATED

The discussion continued:

- Members asked about the ability of the proposed questions to distinguish between instructors and identify problems with course content or teaching. Professor Price answered that, in the pilot study, students indicated that they were better able to differentiate between instructors with the proposed questions and that TLC had worked to eliminate the ambiguity in some of the current questions that were perceived to have two meanings (e.g. Did the instructor speak audibly

and clearly?).

- A member indicated that, in the past,, it was possible to compare results across Faculties as well as Departments, but that the proposed policy did not allow for reports on Faculty distributions.

It was **AGREED** that the text 'Faculty/' would be added to Section 111.3 (F) iii. Section 111.3(F) would read (addition underlined):

"111.3 (F) iii Numerical valued from Tukey's boxplot statistics will be provided to describe the distribution of scores in the Faculty/department:
a) lower cutoff for outlier scores
b) lower hinge (25th percentile)
c) median
d) upper hinge (75th percentile)
e) it is expected that the upper cut-off will always be 5.0 and, therefore, unnecessary to report."

A member asked if it would be possible to gather information regarding whether the course was an elective or not. Professor Price answered that the instructors could include that particular question in their own section of the evaluation.

VOTE ON MOTION The Motion was PUT and **CARRIED.**

56. **FACULTY OF MEDICINE AND DENTISTRY--PROPOSAL FOR THE ADMINISTRATIVE RESTRUCTURING OF THE EXISTING DEPARTMENT OF CELL BIOLOGY AND ANATOMY INTO THE DEPARTMENT OF CELL BIOLOGY AND THE FACULTY DIVISION OF ANATOMY: RECOMMENDATION OF THE GFC ACADEMIC PLANNING COMMITTEE (APC)**

Members considered a proposal, submitted by the Faculty of Medicine and Dentistry, for the (proposed) administrative restructuring of the existing Department of Cell Biology and Anatomy into the Department of Cell Biology and the Faculty Division of Anatomy, to take effect upon the approval of both GFC and the Board of Governors.

It was noted that the GFC Academic Planning Committee had considered and endorsed the above-noted proposal at its meeting of March 24, 1999.

Relevant background information and legislation (Sections 3.1.B and 37, GFC Policy Manual) were before members.

Dr P Gordon, Executive Associate Dean, Faculty of Medicine and Dentistry, was in attendance and presented the proposal.

University of Alberta **Inter-departmental Correspondence**

351 GFC 12/April/1999

to: Ms EA Schoeck
Secretary to GFC

date: March 18, 1999

our file:

from: Mr A Drummond
Secretary, Teaching and Learning Committee
2-5 University Hall

your file:

subject: **Teaching Evaluations: Proposed Revisions to Section 111.3, GFC Policy Manual**

At its meetings of February 9 and March 9, 1999, TLC endorsed, and recommended that GFC Executive endorse and recommend to GFC for its approval, a series of changes to Section 111.3 of the GFC Policy Manual, as laid out in the right-hand column of the attached comparative table.

This recommendation was presented to TLC by its Teaching Evaluation Subcommittee (TES), which was struck in 1997 to study in detail the current policy relating to teaching evaluations and to recommend changes, if any, to that policy. TES met throughout the 1997-98 and 1998-99 academic years, and contracted the Population Research Laboratory to review relevant academic literature on the subject and, resulting from that review, to test a series of new questions in November, 1998.

The literature review and testing strategy were circulated to both students and faculty prior to the test, and the Students' Union, the Graduate Students' Association, and the Association of Academic Staff all commented on the strategy.

The test consisted of 27 questions rated on two different scales. The data from that test were then analyzed extensively by the Population Research Laboratory, and TES recommended that ten questions be selected with a common scale. TLC accepted that recommendation unanimously at its meeting of February 9, after lengthy discussion. TES' report is attached, along with a brief summary of the test results and a description of participating classes.

TLC then asked TES to consider an overhaul of the wording of Section 111.3, and to recommend any other changes that it felt might be required. At its meeting of March 9, TLC approved unanimously, after lengthy discussion, more revisions. Among the changes it recommended was the reporting of a statistical technique known as "Tukey's 'box-and-whisker' plots;" a description of this function is attached for your information.

ION REQUIRED: To consider and approve the changes to n 111.3 of the GFC Policy Manual (Teaching Evaluations) as set in the right-hand column of the attached comparative table.

CUTIVE ACTION: At its meeting of March 29, 1999, Minute f the Executive voted to recommend that GFC approve this proposal.

Relevant background is described in the cover letters.

Andrew Drummond

(For the General Faculties Council Meeting of April 12, 1999)

University of Alberta
Edmonton

Canada T6G 2J9

Office of the
Vice-President (Academic)

3rd Floor, University Hall

Fax (403) 492-1438

March 12, 1999

To: Ellen Schoek
Secretary to GFC

From: Anne Marie Decore
Associate Vice-President (Academic)

You may recall that just over a year ago ~~the~~ when changes to some aspects of the Teaching Evaluation Policy were approved, the GFC Teaching and Learning Committee (TLC) indicated that revisions to the Universal Student Rating of Instruction Questionnaire were underway and would be forwarded in due course. In the ensuing sixteen months, a thorough review of the literature was undertaken and a number of new and revised items and response categories were tested in a pilot study undertaken by the Population Research Centre for TLC. After a very careful analysis of the data from that pilot study, TLC has approved the attached modifications of the GFC Teaching Evaluation policy. I should note that in addition to changes in the Universal Student Rating of Instruction questions, TLC has also undertaken an editorial cleanup of Section 111.3.

I would ask that the proposed revisions to Section 111.3 (Universal Student Ratings of Instruction) of the GFC Policy Manual, as set out in the right-hand column of the attached comparative table, be forwarded to the Academic Planning Committee for information and to the GFC Executive Committee for its consideration. Members of TLC and I would be happy to attend those meetings of APC and the Executive when the revisions are considered.

cc: Dr D R Owram
Mr G Bodnar

111.2 TEACHING EVALUATION	111.2 TEACHING EVALUATION
1. Evaluation of teaching at the University of Alberta serves two purposes: a. Summative - Evaluation provides a review and overview of an instructor's teaching that is an essential element in promotion and tenure decisions. In its summative form, teaching evaluation forms a basis for rewarding excellence, as well as the basis for withholding reward. b. Formative - Evaluation provides helpful feedback to teachers by identifying teaching strengths and weaknesses and, in so doing, giving guidance for the improvement or refinement of teaching skills.	No change.
2. Evaluation of teaching shall be multifaceted. Multifaceted evaluation shall include the Universal Student Ratings of Instruction set out in Section 111.3 and other methods of assessing teaching designed within the individual Faculties to respond to the particular conditions of that Faculty. Such assessments shall include one or more of the following: input from administrators, peers, self, undergraduate and graduate students, and alumni.	No change.
3. Recognizing that the evaluation of teaching at the University shall be multifaceted, Faculty Evaluation Committee (FEC) decisions concerning tenure, promotion or unsatisfactory teaching performance must be based on more than one indicator of the adequacy of teaching.	No change.
4. Assessment of teaching involving input from administrators, peers, self, alumni, or undergraduate and graduate student input additional to the Universal Student Ratings of Instruction should occur annually prior to tenure. For continuing faculty (ie., Categories A1.1, A1.5 and A1.6), such assessment will occur at least triennially.	No change
5. The University shall continue to support University Teaching Services in its education programming which is focused on the development and improvement of teaching and	No change.

learning and its efforts to enhance research in university teaching.	
111.3 UNIVERSAL STUDENT RATINGS OF INSTRUCTION	111.3 UNIVERSAL STUDENT RATINGS OF INSTRUCTION
In recognition of the University's commitment to teaching, the General Faculties Council endorses a ~~universal system of student ratings of instruction. However, the universal rating system~~ is only part of the multi-faceted approach described in Section 111.2.	In recognition of the University's commitment to teaching, the General Faculties Council endorses a system of Universal Student Ratings of Instruction. This system, however, is only one part of the multi-faceted approach described in Section 111.2.
The Universal Student Ratings of Instruction are designed to provide a minimal university-wide base of information on student ratings to the parties listed in this Section. With this purpose in mind, the General Faculties Council adopts the following policies:	No Change.
A. All Faculties shall ensure that evaluation of all instructors and courses shall take place each time a course is offered. The term 'instructors' is meant to include tenured professors, tenure-track professors, sessional instructors, clinical instructors, field supervisors and graduate teaching assistants with responsibilities for courses. The term 'course' is meant to include undergraduate and graduate courses, laboratory courses, non-degree courses, seminars, clinical supervision courses, ~~or~~ reading~~/~~directed study courses. With the exceptions noted in Section 111.3.B, the assessment shall include the Universal Student Ratings of Instruction as set out below.	A. All Faculties shall ensure that evaluation of all instructors and courses shall take place each time a course is offered. The term 'instructors' is meant to include tenured professors, tenure-track professors, sessional instructors, clinical instructors, field supervisors and graduate teaching assistants with responsibilities for courses. The term 'course' is meant to include undergraduate and graduate courses, laboratory courses, non-degree courses, seminars, clinical supervision courses, and reading or directed study courses. With the exceptions noted in Section 111.3.B, the assessment shall include the Universal Student Ratings of Instruction as set out below.
B. ~~The universal student rating system shall not be required in the following situations:~~	B. The Universal Student Ratings of Instruction shall be modified in the following circumstances:
i. ~~courses with fewer than 10 registered students~~;	i. courses with between four and nine registered students shall use a department or Faculty developed questionnaire with non-scored questions, such as: a) comments on the quality of this course; b) suggestions for improving this course; c) comments on the quality of instruction in this course; d) suggestions for improving the instruction in this course.
ii. ~~instructors responsible for fewer than one~~	ii. courses with multiple instructors shall use

~~quarter of the classes in a course~~;	a modified Universal Student Ratings of Instruction questionnaire that will include one set of course-related questions for the entire course and one set of instructor-related questions for each instructor who has taught the equivalent of twenty percent or more of the course. If no instructor is responsible for at least twenty percent of the course, only course-related questions should be used on the questionnaire.
iii. ~~courses where multiple instruction means that no instructor is responsible for one quarter of the course.~~ In ~~such cases,~~ the ~~Dean/Chair~~ shall arrange for an alternate method of ~~assessment~~. Such ~~assessment may~~ include exit interviews with the Chair; questions designed to rate the course rather than multiple instructors; or other means.	iii. in courses with fewer than four registered students or courses such as alternate delivery style courses, the Chair, Director or Dean shall arrange for an alternate method of obtaining student feedback. Such methods could include student course or program exit interviews with the Chair, Director or Dean; or other appropriate means.
C. ~~Summary data from the~~ Universal Student Ratings of Instruction ~~should~~ be reported to both the instructor and the Department Chair/Dean as ~~(1)~~ the number of students responding in each category; ~~(2) the percentage of students responding in each category; and (3)~~ the median score to one decimal point for that question ~~in comparison to the median score for that question in the Department/Faculty~~.	Revised and renumbered as F.
D. Parties having access to ~~results of individual~~ Universal Student Ratings of Instruction shall be the instructor of the course, the department Chair, members of Tenure Committees and Faculty Evaluation Committees. Normally, instructors shall receive the student ratings of instruction within twenty working days after ~~the completion of final examinations but not before~~ the course is complete and the grade sheet has been signed by the Chair, Director or Dean.	Revised and renumbered as G.
The ~~results~~ of the ~~seven~~ universal questions ~~listed below~~ shall be given to the Students' Union and the Graduate Students' Association ~~in full numerical summary form. The Students' Union and the Graduate Students' Association shall make these results available only to~~	Revised

~~University of Alberta students.~~ The Students' Union and Graduate Students' Association acknowledge that the Universal Student Ratings of Instruction are intended only for ~~student~~ use and shall not be made available to other parties. Neither the Students' Union nor the Graduate Students' Association shall undertake further analysis of the data. The results will be made available to students in paper and/or electronic form.	
E. ~~All results given out~~ to students shall have the following cautionary preface: *Student ~~questionnaires form~~ an important part of evaluating teaching effectiveness but cannot be taken alone as a complete assessment of an instructor or course. Factors other than an instructor's teaching ability may influence ratings. These factors include class size, class level, Faculty, time of class, required versus optional course, grade expectations, student GPA, gender, race/ethnicity, age of both students and instructors~~. ~~Finally, note that both professors and students are diverse and that professors who are rated poorly by some students may be judged exceptional by others (and vice versa).~~* Small differences in evaluation should not be considered as meaningful.	Revised and renumbered as H.
~~F.~~ The Universal Student Ratings of Instruction shall take the form of a questionnaire. The following statement of purpose shall be included at the beginning of the questionnaire: *The University of Alberta would appreciate your careful completion of this questionnaire. The results are one important factor in decisions affecting the career of your instructor. The ~~results of the seven~~ questions listed below are available through the Students' Union and the Graduate Students' Association.* *To protect the anonymity of student ~~responses~~ written comments will be typed where the ~~Department~~ Chair/Dean deems it advisable.*	C. The Universal Student Ratings of Instruction shall take the form of a questionnaire. The following statement of purpose shall be included at the beginning of the questionnaire: *The University of Alberta would appreciate your careful completion of this questionnaire. <u>The results help instructors and departments or faculties to initiate constructive change in curriculum and instruction. In addition</u>, the results are one important factor in decisions affecting the career of your instructor. The <u>numerical summaries for the ten</u> questions listed below are available to students through the Students' Union and the Graduate Students' Association.* *To protect the anonymity of students, <u>their</u> written comments will be typed where the Chair, <u>Director or</u> Dean deems it advisable.*

Students who are concerned about the anonymity of their responses should submit their typewritten comments within five working days of the assessment done in class to the Chair ~~of the Department~~, making sure to note the course number, section and name of the instructor.	*Students who are concerned about the anonymity of their responses should submit their typewritten comments within five working days of the assessment done in class to the Chair, Director or Dean, making sure to note the course number, section and name of the instructor.*
Questions about this questionnaire should be addressed to your ~~department~~ Chair or Dean.	*Questions about this questionnaire should be addressed to the Chair, Director or Dean.*
G. The Universal Student Ratings of Instruction shall ~~contain the following questions~~:	D. The Universal Student Ratings of Instruction questionnaire shall use the rating scale Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree to gather responses to the following questions:
~~1. My university year is 1=First 2=Second 3=Third 4=Fourth 5=Post-Degree~~	Delete.
~~2. This course was 1=a requirement 2=an elective 3=other~~	Deleted.
	1. The goals and objectives of the course were clear.
	2. In-class time was used effectively.
	3. I am motivated to learn more about these subject areas.
	4. I increased my knowledge of the subject areas in this course.
	5. Overall the quality of the course content was excellent.
~~3.~~ The instructor spoke ~~audibly and~~ clearly. ~~1. Strongly Disagree 2. Disagree 3. Neutral 4. Agree 5. Strongly Agree~~	6. The instructor spoke clearly.
~~4. The instructor was accessible outside of class. 1. Strongly Disagree 2. Disagree 3. Neutral 4. Agree 5. Strongly Agree 6. Never Tried/Not Applicable~~	Deleted.
	7. The instructor was well prepared.
~~5.~~ The instructor treated the students with	8. The instructor treated the students with

respect. ~~1. Strongly Disagree 2. Disagree 3. Neutral 4. Agree 5. Strongly Agree~~	respect.
	9. The instructor provided constructive feedback throughout this course.
~~6.~~ Overall the instructor ~~is 1. Poor 2. Fair 3. Acceptable 4. Very Good 5.~~ Excellent	10. Overall this instructor was excellent.
~~7. Overall this course is 1. Poor 2. Fair 3. Acceptable 4. Very Good 5. Excellent~~	Renumbered and revised as E. 5.
	These constitute the ten required Universal Student Ratings of Instruction questions. Instructors, departments, and faculties are encouraged to supplement the set of universal questions.
~~These questions shall be followed with a~~ space for comments. The comments will not be released to the Students' Union or the Graduate Students' Association.	The questionnaire shall allow space for comments. Last sentence incorporated in G.
~~H. This policy is designed to provide only a minimal core of questions for students and for cross-university purposes. Faculties, departments and instructors~~ are encouraged to supplement the universal questions. ~~Only the results of the universal questions are made available to students.~~	Revised and incorporated in D. Revised and incorporated in G.
	E. Certain policies are necessary in order to ensure that the Universal Student Ratings of Instruction Questionnaire is administered in as consistent a fashion as possible. These are:
	i. The Universal Student Ratings of Instruction questions and additional instructor, department or Faculty selected questions shall normally be rated in the same class period.
	ii. Questionnaires shall be administered and completed at the beginning of the class period.
	iii. Universal Student Ratings of Instruction shall normally be administered toward the end of the course but not during the last week of classes.
	iv. The instructor shall not distribute the questionnaires; shall not be present in the

	room when the questionnaires are being completed; and shall not collect the questionnaires. Departments or Faculties shall create policies to ensure that other individuals (eg, other instructors, students within the class, teaching assistants) are available to administer the questionnaires. v. The questionnaires shall be taken directly from the class by the person responsible for administration of the questionnaire to the Chair, Director or designate (or, in the case of non-departmentalized Faculties, to the Dean or designate). The Chair or designate shall then transmit the questionnaires for optical scanning and be responsible for transmission of scanned results and comments to the instructor under the conditions set out in ~~the above-noted~~ Section G. F. The numerical summaries for the ten Universal Student Ratings of Instruction questions shall be reported to the Instructor, the Chair, Director or Dean and students. i. the number of students responding in each category; ii. the median score to one decimal point for the question; and iii. numerical values from Tukey's boxplot statistics will be provided to describe the distribution of scores in the department: a) lower cut-off for outlier scores b) lower hinge (25th percentile) c) median d) upper hinge (75th percentile) e) it is expected that the upper cut-off will always be 5.0 and, therefore, unnecessary to report. NOTE: Statistics from Tukey's box-and-whisker plot analysis (John W. Tukey, Exploratory Data Analysis, Addison-Wesley Publishing Company, Inc. 1977) have been selected to describe the distribution of USRI data. These statistics are chosen to achieve two main objectives: (i) summarising skewed data and (ii) identifying outliers from the general population if they exist. The median (middle of a ranked set of numbers) is generally preferred rather than the mean in defining the centre of a skewed data set.

The 25th and 75th percentiles provide information about the spread of individual scores around the median. By definition, half of the scores in a distribution are below the median and 25 percent of the scores are below the 25th percentile. Since this occurs "by definition", these values should not be used to determine whether a particular score is "good" or "bad".

The lower whisker or cut-off, which is 1.5 box lengths below the 25th percentile (box length is the distance from the 25th to the 75th percentile), defines a reasonable limit beyond which any score can be considered an outlier. Outliers are scores that identify ratings of instruction falling outside the usual distribution of the scores for the population being tabulated.

Given the nature of the USRI data, the upper whisker or cut-off (1.5 box lengths above the 75th percentile) will usually be above 5.0, and so need not be reported.

G. Parties having access to numerical summaries of the ten Universal Student Ratings of Instruction questions and student comments shall be the instructor; the Chair, Director or Dean; members of Tenure Committees; and members of Faculty Evaluation Committees.

For questions selected by an instructor, only the instructor shall receive the results. For questions initiated or mandated by a department or Faculty, the results will be reported to the instructor and the Chair, Director or Dean.

Normally, instructors shall receive the results from the student ratings of instruction within twenty working days after the course is complete and the grade sheet has been signed by the Chair, Director or Dean.

Numerical summaries for the ten Universal Student Ratings of Instruction questions shall be given to the Students' Union and the Graduate Students' Association. Results of additional selected questions and student comments shall not be made available to the Students' Union or the Graduate Students' Association.

The Students' Union and Graduate Students' Association acknowledge that the Universal Student Ratings of Instruction are intended only for use by University of Alberta students and

	shall not be made available to other parties. Neither the Students' Union nor the Graduate Students' Association shall undertake further analysis of the data. The results will be made available to students in paper and/or electronic form. The results will not be released in electronic form for at least ten days following the provision of the results to the instructor. H. All results given out to students, Chairs, Directors and Deans shall have the following cautionary preface: *Student questionnaires form an important part of evaluating teaching effectiveness but cannot be taken alone as a complete assessment of an instructor or course. Factors other than an instructor's teaching ability may influence ratings. These factors include class size, class level, Faculty, time of class, required versus optional course, grade expectations, student GPA, gender, race, ethnicity, age of both students and instructors.* *Small differences in evaluation should not be considered meaningful.*
I. The central administration of the University shall undertake the financing of the universal set of questions in support of the University's commitment to teaching.	Renumbered as J.
~~J.~~ Certain policies are necessary in order to ensure that the universal rating system is administered in as consistent a fashion as possible. These are: i. The Universal Student Ratings of Instruction ~~and further mandated department, Faculty or instructor~~ questions shall normally be given on the same day. ii. Questionnaires shall be administered at the beginning of the class period. iii. Universal Student Ratings of Instruction shall normally be administered ~~within the last two weeks of class~~. iv. The instructor shall not distribute the questions; shall not be present in the room when the questionnaires are being	Renumbered as F.

completed; and shall not collect the questions. Departments~~,~~ Faculties shall create policy~~~~ to ensure that other individuals are available to administer the questionnaires~~ eg, other instructors, students within the class, teaching assistants~~.	
v. The questionnaires shall be taken directly from the class by the person responsible for administration of the questionnaire to the ~~department~~ Chair or designate (or, in the case of non-departmentalized Faculties, to the Dean or designate). The Chair or designate shall then transmit the questionnaires for optical scanning and be responsible for transmission of scanned results and comments to the instructor under the conditions set out in ~~the above-noted~~ Section ~~C~~.	
K. Nothing in this section shall prevent instructors from seeking other means of feedback from students during the term.	Renumbered as I.

TEACHING EVALUATION SUB-COMMITTEE REPORT

Preamble:

The Teaching Evaluation sub-Committee has conducted a careful review of methodologies and questions for use in Universal Student Rating of Instruction at the University of Alberta. This has included a literature review and campus wide "field test" of a proposed instrument. Following detailed statistical analysis by the Population Research Laboratory and by Dan Precht of Computing & Network Services the Teaching Evaluation sub-Committee submits the following proposals:

1. Our field test showed no advantage for using a 5 point, time-based scale (Hardly Ever, Occasionally, Sometimes, Frequently, Almost Always), so we recommend that we retain the current scale except that it should use only one set of descriptors, namely:

   Strongly Disagree; Disagree; Neutral; Agree; Strongly Agree

2. Our field test used 27 new questions; the statistical analysis showed that a number of these could be dropped without serious loss. In reviewing the questions to be dropped and those to be retained the sub-committee was particularly careful to avoid high inference questions, and to ensure that four areas, which were identified as being important to the students through the field test, would be adequately addressed. These four areas were:

   Course Quality, Instructional Quality, Communication, and Course Organization.

   The following questions (some of which address more than one area on the list) are recommended by the sub-Committee:

   1. The goals and objectives of the course were clear.
   2. In-class time was used effectively.
   3. I am motivated to learn more about these subject areas.
   4. I increased my knowledge of the subject areas in this course.
   5. Overall, the quality of the course content was excellent.
   6. The instructor spoke clearly.
   7. The instructor was well prepared.
   8. The instructor treated students with respect.
   9. The instructor provided constructive feedback throughout this course.
   10. Overall, this instructor was excellent.

3. We feel that at this stage it would be premature to make any recommendations regarding the use of this instrument, but we think it important that users (students, instructors, chairs and FECs) should have a clear understanding of its purposes, strengths, and limitations. If it is to become a functional part of a University wide effort to identify and reward instructional success, it has to have the respect of all parties.

4. The sub-Committee began discussing the collection of demographic data with each questionnaire, for the purpose of future research, but has no recommendation at this stage. The discussion is incomplete, but this topic should be revisited if the above proposals are accepted.

Submitted on behalf of the sub-Committee

Mick Price
1999 February 9

Note from the University Secretariat: At its meetings of February 9, 1999 and March 9, 1999, the GFC Teaching and Learning Committee (TLC) considered the attached material in its discussions regarding the proposed revisions to Section 111.3 (Universal Student Ratings of Instruction). It has been attached for your reference.

TESTING OF NEW ITEMS FOR THE UNIVERSAL STUDENT RATINGS OF INSTRUCTION

BRIEF SUMMARY OF STATISTICAL RESULTS 1999.01.19

- A total of 4,546 test forms were completed by undergraduate students at the University of Alberta for the purpose of field testing new items for the Universal Student Ratings of Instruction.

- Classified by class size, 18.5% of the classes in which the testing occurred had enrollments of 35 students or less, 33.1% of the classes had enrollments between 36 and 99 students, and 48.4% had enrolments of 100 students or more.

- Classified by course level, 31.2% of the courses in which the testing occurred were 100 level courses, 26.9% of the courses were 200-level courses, 26.0% were 300 level courses, and 16.0% were 400 level courses.

- Classified by type of faculty, 17.2% of the courses in which testing occurred were offered by the Faculty of Arts, 34.4% of the courses were offered by the Faculty of Science or the Faculty of Engineering, 21.3% of the courses were offered in one of the five health sciences faculties, and 27.0% of the courses were offered in other faculties, except for the Faculté Saint-Jean

- Overall, ratings of instruction ranged across all possible response categories for all items tested.

- Overall, the instructors evaluated using the new USRI test items typically received positive evaluations by the students who participated in testing the field items. The modal response given by students using a five-point scale is to select the number "4". For example, the instructors of individual courses are most often rated at "very good" by students.

- Asked to indicate which set of questions (old or new) they preferred to use as a basis for evaluation of the quality of teaching at the University of Alberta, the majority (70.7%) of the 4207 responses to this question indicated that students preferred the new set of USRI items. This finding was consistent across course levels, faculties, and class sizes.

- In general, responses to individual USRI items were positively and significantly related to each other. There was a high level of internal consistency among the test items.

Description of Participating Classes

		Number Responding			
		N of Classes	% of Classes	N of Students	% of Students
Discipline	Arts	26	25.0%	792	17.2%
	Science-Engineering	31	29.8%	1565	33.9%
	Health Related	19	18.3%	979	21.2%
	Other	28	26.9%	1278	27.7%
Course Level	100 level	27	26.0%	1422	30.8%
	200 level	27	26.0%	1237	26.8%
	300 level	26	25.0%	1186	25.7%
	400 level	24	23.1%	769	16.7%
Grouped Class Size	<35	48	46.2%	844	18.3%
	35-99	35	33.7%	1558	33.8%
	>99	21	20.2%	2212	47.9%
Total		104	100.0%	4614	100.0%

Section 1.2

Describing Distributions with numbers.

Two of the most important features of the distribution of a quantitative variable can be described using numerical measures,

- Its center, and
- The spread of its values about the center

Two of the numbers we use to describe the center of a distrubution (i.e., the location where roughly half the values are below it and the other half above it) are:

1. The mean
2. The median

-- The mean of the distribution of a quantitative variable is the arithmetic average of its values.

$$\bar{x} = \frac{\sum x_i}{n} = \frac{x_1 + x_2 + x_3 + \cdots x_n}{n}$$

--The median is the "middle value". It is located after all the observations have been put in order.

- --arrange all observations in increasing (or ascending) order
- --if n is odd, then the median is the (n+1)/2 th observation
- --if n is even, the median is the mean of the two center observations

Comparing the mean and the median.

-- Because the mean is the arithmetic average of all the values in a set of data, it is strongly influenced by any extreme observations that are included in the set. The mean always missrepresents (either underestimates or overastimates) the center of distributions that are skewed either to the left or to the right.

-- The median on the other hand is resistant to any extreme observations that the data set may include. It is always a better choice to use the median to describe the location of the center for skewed distributions.

-- For symmetric distributions the mean and median should both be fairly close (or even equal) to each other.

Examples 1.7 & 1.8 (pages 37 & 39) illustrate these differences. It is worth mentioning that both the mean and the median give valuable information about the distribution of a quantitative variable. As it is pointed out in the book, the median in Example 1.7 should be used to describe the number of hysterectomies performed by a "typical" male doctor. On the other hand, the mean is a good measure for the total number of hysterectomies performed by all the male odctors.

Measuring spread

The mean and median alone do not describe the distribution of a variable completely. Numerical measures of spread give an idea of the variability in the values of a variable.

Common measures of spread

- Range = max obs-min obs
- Quartiles

-- list the observations in increasing order
-- the first (lower) quartile is the median of the first half of the data (Q_1)
-- the second quartile is the median
-- the third (upper) quartile is the median of the second half of the data (Q_3)

- Inter Quartile Range (IQR)

-- IQR = Q_3 - Q_1

- Five Number Summary

Minimum, Q_1, M, Q_3, Maximum

- The Boxplot

--The five numbers - (2) extremes, (2) quartiles, and the median tell us a great deal about a set of data. These five numbers are also used to draw a different kind of plot, the BOX PLOT.

Draw a box plot as follows:

- Step1: Find the five-number-summary.
- Step2: Mark the locations of the median, quartiles, and extremes below a number line.
- Step3: Draw a box between the two quartiles. Mark the median with a line across the box. Draw two "whiskers" from the quartiles to the extremes.

- -- Outliers:
  Data values that are substantially larger or smaller than the other values are refered to as outliers.

The 1.5 x IQR rule for outliers:
Observations that fall below Q_1 - 1.5 x (IQR), or above Q_3 + 1.5 x (IQR) are, according to this rule, identified as potential outliers.

Note:
When outliers are present in the data, then a modified boxplot must be drawn. Draw a modified boxplot by ending the whiskers at the most extreme observations still within 1.5 x IQR of the quartiles and plot all of the outliers individually.

Example 1.

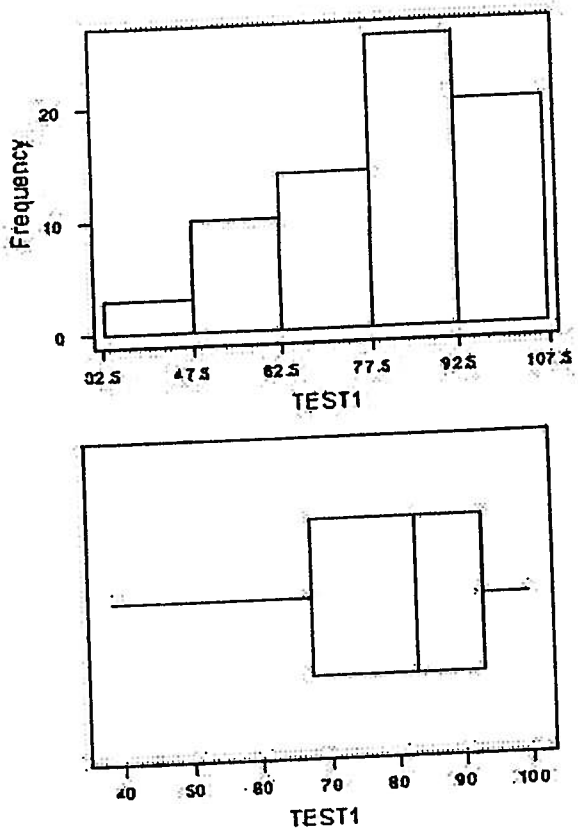
The following is a numerical summary of exam scores on TEST 1 from a previous Math 115 class.

Variable	N	Mean	Median	Min	Max	Q_1	Q_3

tributions

Variable	N	Mean	Median	Min	Max	Q_1	Q_3
TEST1	73	79.21	83.00	38.00	100.00	67.50	93.00

A box plot and a histogram for these results is shown below



2. Variance and The standard deviation.

The variance of the n observations $x_1, x_2, \ldots, x_n$ with mean $\bar{x}$ is defined as:

$$s^2 = \frac{1}{n-1}\sum (x_i - \bar{x})^2$$

The standard deviation of the n observations $x_1, x_2, \ldots, x_n$ with mean $\bar{x}$ is defined as:

$$s = \sqrt{\frac{1}{n-1}\sum (x_i - \bar{x})^2}$$

Properties of s:

- -- Measures the spread around the mean
- -- Should be used with the mean not the median
- -- if s = 0, then all the observations have the same value
- -- The larger the s the more spread out the data is
- -- s is strongly influenced by outliers

NOTE:

– For describing skewed distributions the five-number summary is preferred.
– For describing symmetric distributions the mean and standard deviation are preferred.

**An alternative formula for computing the variance which is easier to use is:

$$s^2 = \frac{1}{n-1}\left[\sum x_i^2 - \frac{1}{n}\left(\sum x_i\right)^2\right]$$

NOTE: I recommend that you use a calculator to compute the mean and the standard deviation for a set of data. It could be educational though to see how the formulas are used to compute these numbers also. Click here to see an example.

Boxplots

The *five-number summary* is an abbreviated way to describe a sample. The five number summary is a list of the following numbers:

1. Minimum
2. First (Lower) Quartile, Q_1
3. Median, $\tilde{X}$
4. Third (Upper) Quartile, Q_3
5. Maximum

The five number summary leads to a graphical representation of a distribution called the *boxplot*. Boxplots are ideal for comparing two nearly-continuous variables. To draw a boxplot (see the example in the figure below), follow these simple steps:

1. The ends of the box (hinges) are at the quartiles, so that the length of the box is the IQR .
2. The median is marked by a line within the box.
3. The two vertical lines (called *whiskers*) outside the box extend to the smallest and largest observations within $1.5 \times \mathrm{IQR}$ of the quartiles.
4. Observations that fall outside of $3 \times \mathrm{IQR}$ are called *extreme outliers* and are marked, for example, with an open circle. Observations between $1.5 \times \mathrm{IQR}$ and $3 \times \mathrm{IQR}$ are called *mild outliers* and are distinguished by a different mark, e.g., a closed circle.

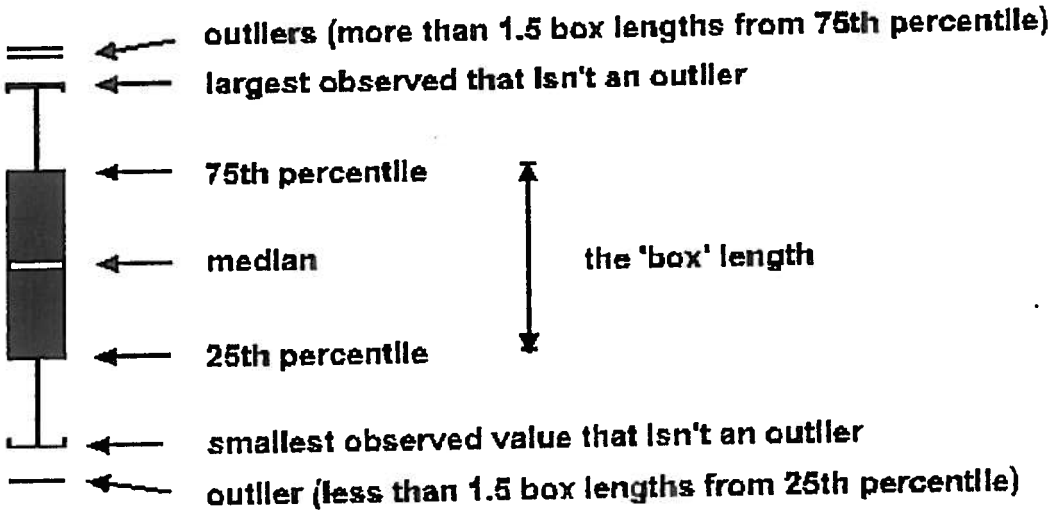
EXAMPLE: To illustrate boxplots, the figure below puts boxplots side by side of the same four data sets that had histograms in the figure in Week 1.

Jan Lethen
Wed Nov 13 16:20:46 CST 1996

Understanding Boxplots

[return to main document]

A boxplot is a display that summarizes information about data contained in a histogram. Rather than plotting the actual values, a boxplot displays summary statistics of the distribution. It indicates the median, the 25th percentile, the 75th percentile, and values that are far removed from the rest (outliers) as shown schematically in the figure.



The white line in the center of the 'box' (the black filled region) indicates the median of the distribution. The horizontal lines that form the boundary of the box indicate the 25th percentile (lower boundary) and the 75th percentile (upper boundary). Fifty percent of the cases fall within this box.

The boxplot includes cases with outlying values defined as a case with a value that is more than 1.5 box lengths from the upper or lower edge of the box.

Also included are the largest and smallest observed values that aren't outliers. Lines are drawn from the ends of the box to these values (sometimes called whiskers).

[return to main document]

Last modified 03-Jan-97
gauthier@phy.duke.edu

Why 1.5?

This exercise or assessment item for high school students nicely links boxplots and outliers with the normal distribution.

*****

Many students are curious about the '1.5*IQR Rule', i.e. why do we use Q1 - 1.5*IQR (or Q3 + 1.5*IQR) as the value for deciding if a data value is classified as an outlier? Paul Velleman, a statistician at Cornell University, was a student of John Tukey, who invented the boxplot and the 1.5*IQR Rule. When he asked Tukey, 'Why 1.5?', Tukey answered, 'Because 1 is too small and 2 is too large.'

It has been shown that this is a reasonable rule for determining if a point is an outlier, for a variety of distributions. This question asks the student to demonstrate this for the normal distribution.

1. Assuming that a dataset is approximately normally distributed, show that about 1 data value in 100 would be classified as outliers, using the 1.5*IQR rule.

2. What percentage of data values would be classified as outliers if we adopt a 2.0*IQR rule instead?

Worked Solution

a. Assume a standard normal distribution. Let the z-value for the1st quartile be z_1 and the z-value for the 3rd quartile be z_2, ie $P(x < z_1) = 0.25$ and $P(x > z_2) = 0.25$

From a normal distribution table, $z_1 = -0.674$ and $z_2 = 0.674$. Hence the IQR is $z_2 - z_1 = 1.348$.

The z-score which is 1.5*IQR below the first quartile is $z_3 = -0.674 - 1.5*1.348 = -2.596$, while the z-score which is 1.5*IQR above the third quartile is $z_4 = 2.596$

From the standard normal table, $P(x < z_3) = 0.5 - 0.4953 = 0.0047$, while $P(x > z_4) = 0.0047$

Hence the probability that a value is an outlier is $2 * 0.0047 = 0.0094$ which is approximately 0.01. Therefore about 1 data value in 100 would be an outlier if the distribution was Gaussian.

b. The first part of the analysis is the same, ie $z_1 = -0.674$ and $z_2 = 0.674$. Hence the IQR is $z_2 - z_1 = 1.348$.

$z_3 = -0.674 - 2.0*1.348 = -3.37$, while $z_4 = 3.37$.

From the standard normal table, $P(x < z_3) = 0.5 - 0.4996 = 0.0004$, while $P(x > z_4) = 0.0004$

Hence the probability that a value is an outlier is $2 * 0.0004 = 0.0008$; hence less than 1 data value in 1000 would be classified as an outlier using the 2.0*IQR rule.

Tukey Box-and-Whisker Plot

Question 15: Overall, the instructor was