Student Evaluations of Teaching

Task Force Final Report

Respectfully submitted to the Office of the Provost on July 19, 2012

Report Outline

I. Executive Summary
II. Task Force Overview
    a. Charge
    b. Membership
III. Subcommittees
    a. Membership
    b. Questions Explored
IV. Item Subcommittee Report
V. Data Analysis Subcommittee Report
VI. *Revised* Implementation Subcommittee Report
VII. Evaluating Teaching Subcommittee Report
VIII. Default Evaluation Items
IX. Evaluation Item Pool

I. Executive Summary

Carol A. Hurney, Ph.D.

The Student Evaluations of Teaching Task Force recommends that the academic administration of James Madison University set up a standing Course Evaluation Committee to work with the Blue Project Implementation team to provide guidance and support for the successful migration of all JMU course evaluations, in their current state, to the Blue on-line system by Fall 2015. The task force also recommends that academic units model best practices when analyzing course evaluation results. Most importantly, the task force strongly encourages academic units to implement other strategies to evaluate teaching to complement and contextualize results obtained from student evaluations of teaching.

The course evaluation item committee recommends that each academic unit migrate their unique evaluation instrument to the online system. To support instrument development, this sub-committee compiled an extensive, comprehensive set of quantitative and qualitative items to add to the item pool in the Blue system. These items can be added to unit-level instruments or selected by faculty to customize instruments to provide formative feedback on discrete aspects of the teaching and learning environment. The item sub-committee also compiled a list of best practices that academic units can use when writing new items or revising current items. Although, the item subcommittee compiled a comprehensive set of items for a default instrument, they do not recommend the implementation of university-wide items. Rather, they suggest that the default instrument can be used to guide units that have not developed a robust instrument or allow other units guidance on revising current instruments. Finally, the item sub-committee provides insights from the literature regarding best practices in item writing and instrument design.

Some members of the task force agree that each academic unit should implement independent evaluation tools/items, while other members of the task force support the implementation of select university-wide items on all course evaluations. Differing opinions regarding the value and use of university-wide items reflect concerns with data use/access and student messaging. Many task force members expressed concerns that online course evaluation data is not secure and could potentially be accessed and misused by various individuals and groups. Although the Blue system is only accessible via secure, JMU-authenticated connections, task force members expressed concerns regarding access to this information via avenues such as the Freedom of Information Act. Many members of the task force also expressed strong concerns regarding the use and potential misuse of data from university-wide items. In particular, there is concern that results from these items could be used to compare faculty or generate lists of faculty scores which could result in the misuse these data in decisions regarding promotion, tenure, and merit pay allocations Support by some members of the task force for the use of university-wide items reflect the impact these items could have on student messaging and instrument analysis. University-wide items could reinforce essential aspects of teaching that the entire JMU academic community values. Data from these items could also be used for

psychometric analysis to determine instrument reliability and validity. Currently, the task force is not aware of any academic units at JMU that measure the reliability and validity of the instruments deployed to evaluate teaching.

The data analysis sub-committee recommends that academic units follow scholarly practices when evaluating course evaluation data. In particular, data analysis should reflect response averages, distribution of responses and standard deviations. The Blue system reports data in this fashion, however it unclear at this moment how much flexibility units will have in designing report outputs from the system. When interpreting evaluation results, the task force recommends that units avoid comparing averages in favor of viewing results longitudinally (sample reports provided). Members of the task force expressed concern that administrators could too easily access results of individual faculty members stored in the Blue system. Thus, the task force recommends that only faculty and academic unit heads have direct access via the Blue system to individual faculty reports. JMU administrators would need to contact the academic unit to view or discuss reports of individual faculty when necessary. Only aggregate reports (samples provided) should be directly available via the Blue system to members of the upper administration. Although, the lack of common university-level items will severely limit the use of aggregate reports for the comprehensive analysis of the teaching landscape at JMU.

The implementation sub-committee, in consultation with the Blue Project Implementation, team developed a robust implementation plan. This plan was revised after the Summer 2012 pilot to reflect increased understanding of the functionalities of the Blue system, time for the Course Evaluation Committee to form and begin work with the academic community and the Blue Project Team, and opportunities to work directly with academic unit heads to increase awareness of the functionalities and processes related to successful use of the Blue system. Other key elements of the implementation process reflect practices utilized by other campuses, such as Virginia Tech, and include multiple opportunities to engage all members of the academic community, most notably faculty, in the implementation process. Transparency regarding implementation, in addition to clear lines of communication between the administration and the faculty, will help ensure the successful development of policies and procedures for the successful implementation and use of the Blue system at JMU.

Finally, the evaluating teaching sub-committee recognizes that teaching is a complex, multi-dimensional practice, which cannot be evaluated effectively by any single instrument. Robust and reflective strategies that provide meaningful information to improve the quality of student experience and of instructors' professional lives could be strengthened significantly by a thoughtful, holistic approach to the evaluation of teaching. Academic units should explore the array of strategies recommended by this subcommittee. These tools span various dimensions of teaching, and offer varied opportunities to enhance the formative feedback provided to faculty. Given that overwhelming majority of instructors at JMU are committed and conscientious practitioners of their craft who welcome constructive feedback regarding their professional growth, additional strategies will help contextualize the course evaluation results by providing guidance to instructors and departments engaged in the process of developing an evaluation strategy for teaching.

II. Task Force Overview

a. Task Force Charge

Develop a contextualized course evaluation instrument that will be deployed online.

Issues to be addressed by task force in achieving charge:

- awareness of current literature in this area
- respecting the summative and formative roles of course evaluations
- defining what information we want to gather from course evaluations
- item development: tiered questions for division, college, department, course,
- psychometric testing of items
- issues and best practices regarding on-line course evaluations
- presenting best practices on using the information

b. Task Force Membership

- ***Chair – Carol A. Hurney, Center for Faculty Innovation, Biology***
- *Christie Liu, Center for Instructional Technology*
- *Keston Fulcher Center for Assessment and Research Studies*
- *Mark Parker, English*
- *Joanna Mott, Biology*
- *Catherine Crummett, University Studies*
- *Brian Utter, Physics*
- *Christopher Fox, Computer Science*
- *Joanne Doyle, Economics*
- *Alison Sandman, History*
- *Jonathan Gibson, Music*
- *Margaret Kyger, Exceptional Education*
- *Sam Prins, Mathematics and Statistics*
- *Adebayo Ogundipe, Engineering*
- *Bill O'Meara, Philosophy & Religion*
- *Morgan Benton, Integrated Science & Technology*
- *Andreas Broscheid, Political Science*

III. Subcommittees

a. Membership

Item	Bill O'Meara Chris Fox Jon Gibson Joanna Mott
Data Analysis	Samantha Prins Keston Fulcher Alison Sandman Brian Utter Carol Hurney
Implementation	Christie Liu Catherine Crummett Andreas Broscheid Kristin Sowden
Evaluation Teaching	Morgan Benton Mark Parker Adebayo Ogundipe Maggie Kyger

b. Questions Explored

I. SET Items
   a. What is the value, if any, of having common questions across divisions and college?
   b. What questions are best used for summative and/or formative assessment purposes?

II. Data Analysis
   a. How should units format or organize SET data?
   b. What is the role of qualitative questions in SET data analysis?
   c. What are the best practices in analyzing SET data?
   d. What kind of guidance should we offer to units with respect to data analysis

III. Implementation
   a. How should the on-line system be implemented at JMU?
   b. How should the use of the products of this Task Force be implemented at JMU?
   c. How can we benefit from best practices at JMU regarding SET use and analysis?

IV. Evaluating Teaching
   a. What are other ways to evaluate teaching?
   b. What are the best ways to inform faculty, units, and PACs regarding evidence-based strategies to evaluate teaching?

IV. Item Sub-Committee Report

Members

- **Christopher Fox**
- **Jonathan Gibson**
- **Joanna Mott**
- **Bill O'Meara**

The Items/Literature Sub-Committee focused its efforts on the four following tasks:

- reviewing literature pertaining to the construction of evaluation instruments
- constructing a pool of SET items, from which the University, colleges, departments, and individual instructors might construct evaluation instruments for online deployment
- devising a concise set of guidelines, shaped by SET literature and "best practices," for academic units and individuals wishing to write their own items for incorporation into their evaluation instruments
- recommending practices pertaining to the construction of SET instruments at JMU

The results of our study of these four areas are as follow:

I. The SET literature as a whole is dominated by questions of effectiveness, limitations, implementation, and faculty/student resistance. The literature pertaining specifically to construction of evaluation instruments constitutes a small but considerable sub-set of SET research. This literature has shaped the construction of our item pool, and in particular our guidelines for writing items.

II. The SET Item Pool is attached. The Pool is extensive, with the thought of providing ample guidance primarily on the Academic Unit level. The Items/Literature Sub-Committee has chosen to err on the side of inclusion, even where multiple items were only subtly different in nature or aim.

III. A brief set of "Guidelines for Constructing Your Own Questions" appears at the end of the SET Item Pool. Here, concision was paramount. We attempted to distill a number of "best practices" into a package that would be easily digestible by all faculty members.

IV. As a result of careful consideration and lengthy discussion, the Items/Literature Sub-Committee recommends the following:

Items on the evaluation instrument should be selected on the Academic Unit and individual instructor level, and possibly also on the College level. We do not recommend the inclusion of any University-wide items, for the following reasons: 1) There seems to be no considerable, demonstrable need for University-wide items. 2) Such a policy protects

the autonomy of individual departments and greatly increases the chance of faculty and departmental "buy in," which is particularly important given the inevitable resistance to the migration to a University-wide online delivery system. However...

We recommend that the online evaluation system be populated initially with a default evaluation instrument (Recommended Default Evaluation attached), only after which Academic Units and Colleges could alter, remove, or add items. This default evaluation would serve as a positive example, and would also guard against the possibility of a missing course evaluation, should any given Academic Unit fail to input their items by the deadline.

We recommend that individual faculty members be permitted and encouraged to add items to their course evaluations, whether these items are chosen from the Item Pool or constructed by the faculty member. Furthermore, we recommend that any items added by the individual faculty member not be included in the standard reports made available to AUHs, PACs, et al., unless the evaluated faculty member requests their inclusion.

References

- Berkeley "Tools for Teaching" SET Guidance, Materials
  http://teaching.berkeley.edu/bgd/ratingforms.html

- University of Toronto Review of Student Evaluations of Teaching
  http://senate.dal.ca/Files/meeting_package/Student_Course_Evaluations_(Higher_Ed_Quality_Council_of_Ont.pdf

V. Data Analysis Subcommittee Report

Members

- **Keston Fulcher**
- **Carol Hurney**
- **Samantha Prins**
- **Alison Sandman**
- **Brian Utter**

Best Practices For Interpreting Quantitative Response Items

1. We suggest that instructors and unit heads use the
    a. average response in conjunction with the standard deviation of responses, and
    b. the distribution of responses. Examples: percent of responses at each Likert scale level; percent of students responding at one or both extremes of the scale e.g. % responding agree or strongly agree.

    Users are encouraged to interpret the distribution of responses rather than simply comparing averages, which don't adequately reflect the range of responses (for example, a majority of neutral ratings versus a large number of low and high ratings).
2. Factors that may influence SET ratings are size of class, style of teaching, new or newly revised versus established courses, introductory for non-majors versus higher division for majors, expected student workload, level of course, and discipline (Franklin 2001; Algozzine *et. al.* 2004; Aleamoni 1999; Benton and Cashin 2012). As such, caution should be used in comparing SET ratings across courses, departments, colleges or other units. Comparison across multiple sections of a single course may be appropriate as long as the item(s) being compared are relevant to each course and identically worded with the same Likert scale. If norming (comparative rating) is to be done, we recommend methods along the lines of Neumann (2000), in which courses are compared to similar courses to minimize such differences.
3. SET responses for individual faculty are best viewed longitudinally, with previous semesters or comparable courses providing context for the current semester or course.
4. SET responses from small classes (number of responses below 15-20) or classes with low response rate (below 66% - 75%) are unreliable. (Felder and Brent 2008; Hobson and Talbot 2001; Benton and Cashin 2012).
5. Small differences in average SET ratings should not be used to separate faculty for merit or promotion purposes. Instead, since reliability of SETs is insufficient to make fine distinctions, it is recommended that broad categories be used when comparing faculty performance. For example, ratings could be categorized as *need for improvement*, *similar to most faculty*, and *very good* (Neumann 2000). We realize that this suggestion does not totally alleviate the problem because some scores will undoubtedly lie on the cusp of two categories.
6. If you would like to assess particular aspects of your course that have not been addressed by your SET report, then you can add additional questions to the SET

instrument for the next offering of this course. When adding items, be mindful of keeping the total number of SET questions to at most 15-20.

Best Practices For Interpreting Qualitative Response Items

1. Look for trends in responses rather than anomalies. The comments that you should pay attention to are those that occur frequently in a single class or that occur across multiple classes.
2. Choose themes based on learning objectives of the course or other concerns e.g. physical constraints of the room, critical thinking, prior to reading responses and group responses into these themes.
3. Use the online SET software to view qualitative responses at particular levels of an associated quantitative question e.g. comments on the course overall by students who rated the course overall as excellent versus poor.

References

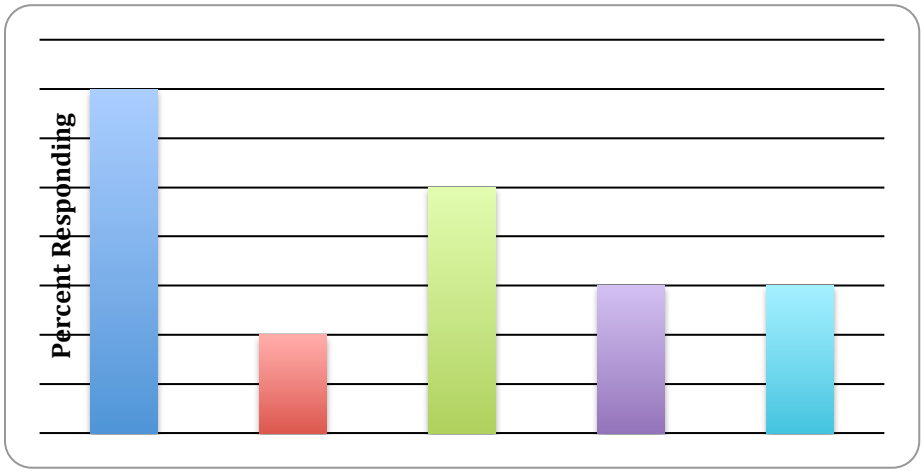
- Lawrence M. Aleamoni. "Student Rating Myths Versus Research Facts from 1924 to 1998." *Journal of Personnel Evaluation in Education* 13, no. 2 (1999): 153-166.
- Bob Algozzine, John Beattie, Marty Bray, Claudia Flowers, John Gretes, Lisa Howley, Ganesh Mohanty, Fred Spooner. "Student Evaluation of College Teaching." *College Teaching* 52, no. 4 (2004): 134-141.
- Benton, Stephen L. and Cashin, William E. *Student Ratings of Teaching: A Summary of Research and Literature.* Idea Paper #50. The Idea Center, 2012.
- Felder, Richard M. and Brent, Rebecca. "Student Ratings Of Teaching: Myths, Facts, And Good Practices." *Chemical Engineering Education* 42, no. 1 (2008): 33-34.
- Franklin, Jennifer. "Interpreting the Numbers: Using a Narrative to Help Others Read Student Evaluations of Your Teaching Accurately." *New Directions for Teaching and Learning* 87, (2001): 85-100.
- Hobson, Suzanne M. and Talbot, Donna M. "Understanding Student Evaluations: What All Faculty Should Know." *College Teaching* 49, no. 1 (2001): 26-31.
- Neumann, Ruth. "Communicating Student Evaluation of Teaching Results: Rating Interpretation Guides (RIGs)." *Assessment and Evaluation in Higher Education* 25, no. 2 (2000): 121-134.

SAMPLE INDIVIDUAL FACULTY REPORT FOR A SINGLE COURSE

The goal of such a report is to summarize the responses of students to a single instance of a course.

Math 220: Elementary Statistics – Section 1
Instructor: John Brown
Semester: Spring 2012
Number of students completing SET: 100

Item 1: The instructor expressed clear expectations for my learning and performance



No. Responses: 100
Mean: 2.65

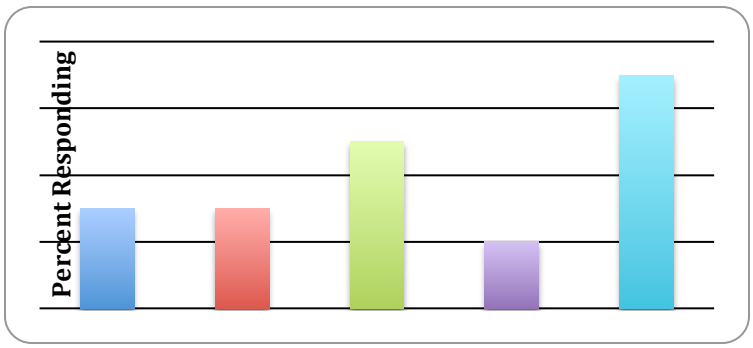
% Agree or Strongly Agree: 30
% Disagree or Strongly Disagree: 45

Item 2: What characteristics of the instructor were most valuable to your learning?

- Response 1 in full
- Response 2 in full
- Response 3 in full
- …
- Response 100 in full

No. Responses: 100
Mean: 3.35

Item 3: Overall, the instructor was an effective teacher.



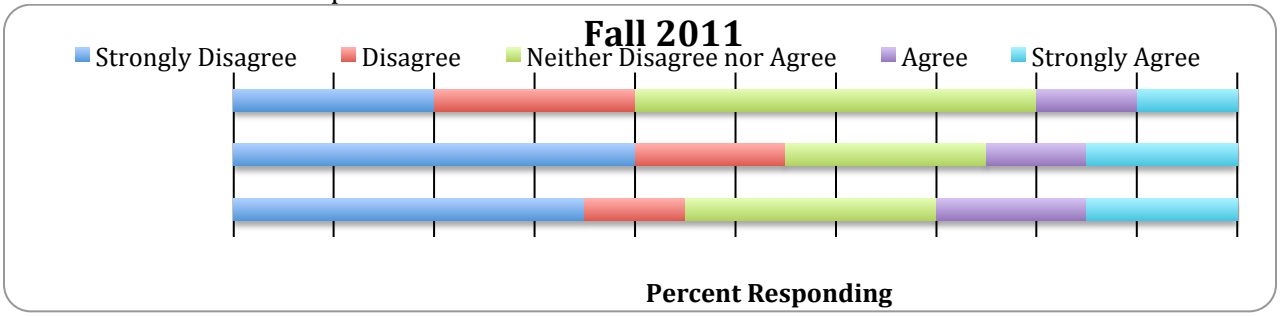
% Agree or Strongly Agree: 30
% Disagree or Strongly Disagree: 45

SAMPLE INDIVIDUAL FACULTY REPORT FOR A SINGLE SEMESTER - VERSION A

The goal of such a report is to compare different courses taught by the same instructor. This comparison should be done with caution. Items selected for comparison should be equally appropriate to each course compared and should be identically worded with the same Likert scale used in each.

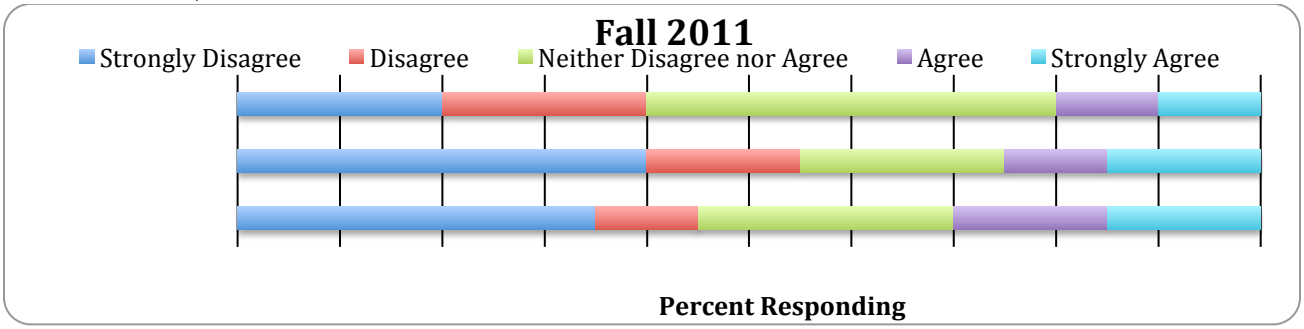
Fall 2011
Instructor: John Brown
Courses: Math 220-001, Math 220-002, Math 280

Item 1: The instructor respects students.



Fall 2011	No. Responses	Average	Std. Dev.	%Agree or Strongly Agree	%Disagree or Strongly Disagree
Math 220-001	34	2.65	1.47	30	45
Math 220-002	34	2.45	1.47	25	55
Math 280	22	2.7	1.19	20	40

Item 3: Overall, the instructor was an effective teacher.

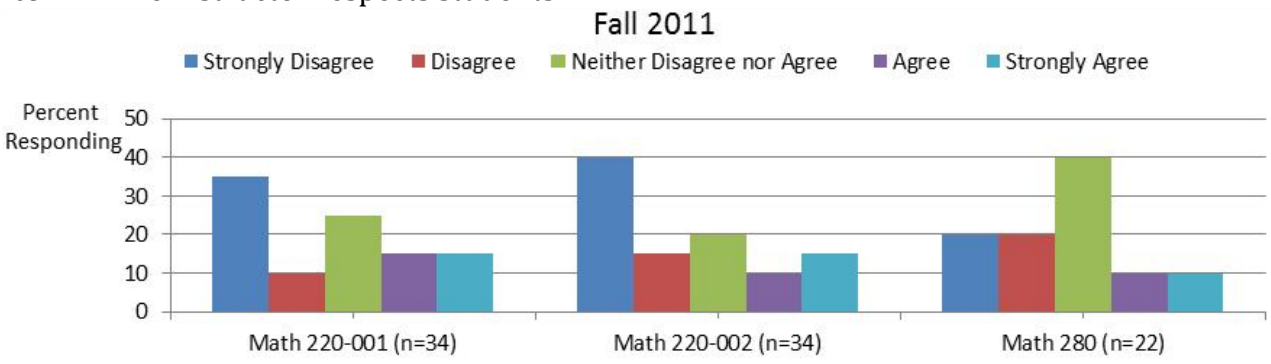


Fall 2011	No. Responses	Average	Std. Dev.	%Agree or Strongly Agree	%Disagree or Strongly Disagree
Math 220-001	34	2.65	1.47	30	45
Math 220-002	34	2.45	1.47	25	55
Math 280	22	2.7	1.19	20	40

SAMPLE INDIVIDUAL FACULTY REPORT FOR A SINGLE SEMESTER - VERSION B

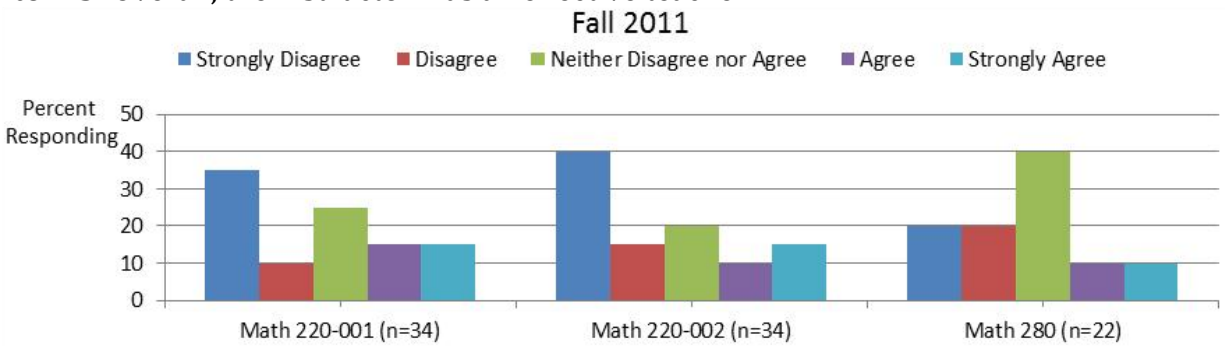
The goal of such a report is to compare different courses taught by the same instructor. This comparison should be done with caution. Items selected for comparison should be equally appropriate to each course compared and should be identically worded with the same Likert scale used in each. The versions illustrate graphics options that we would like available.

Fall 2011
Instructor: John Brown
Courses: Math 220-001, Math 220-002, Math 280
Item 1: The instructor respects students.



Fall 2011	No. Responses	Average	Std. Dev.	%Agree or Strongly Agree	%Disagree or Strongly Disagree
Math 220-001	34	2.65	1.47	30	45
Math 220-002	34	2.45	1.47	25	55
Math 280	22	2.7	1.19	20	40

Item 3: Overall, the instructor was an effective teacher.



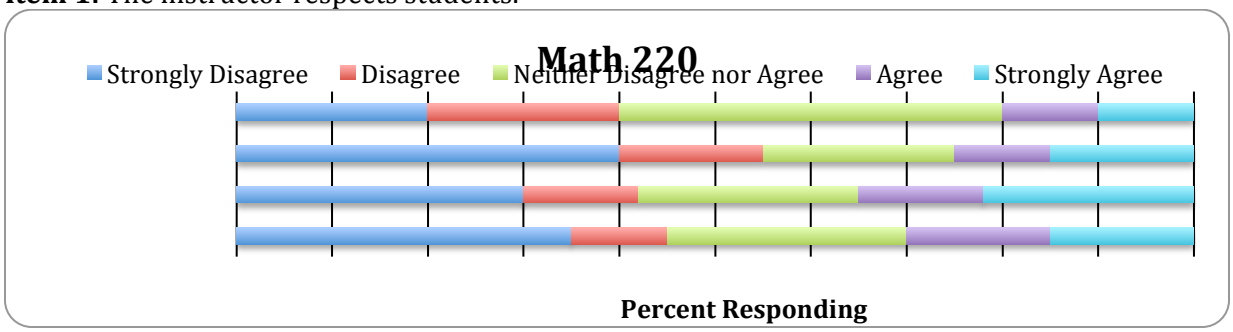
Fall 2011	No. Responses	Average	Std. Dev.	%Agree or Strongly Agree	%Disagree or Strongly Disagree
Math 220-001	34	2.65	1.47	30	45
Math 220-002	34	2.45	1.47	25	55
Math 280	22	2.7	1.19	20	40

SAMPLE LONGITUDINAL INDIVIDUAL FACULTY REPORT FOR A SINGLE COURSE

The goal of such a report is to compare different sections of the same course taught by the same instructor. Items selected for comparison should be equally appropriate to each section and should be identically worded with the same Likert scale used in each. The versions illustrate graphics options that we would like available.

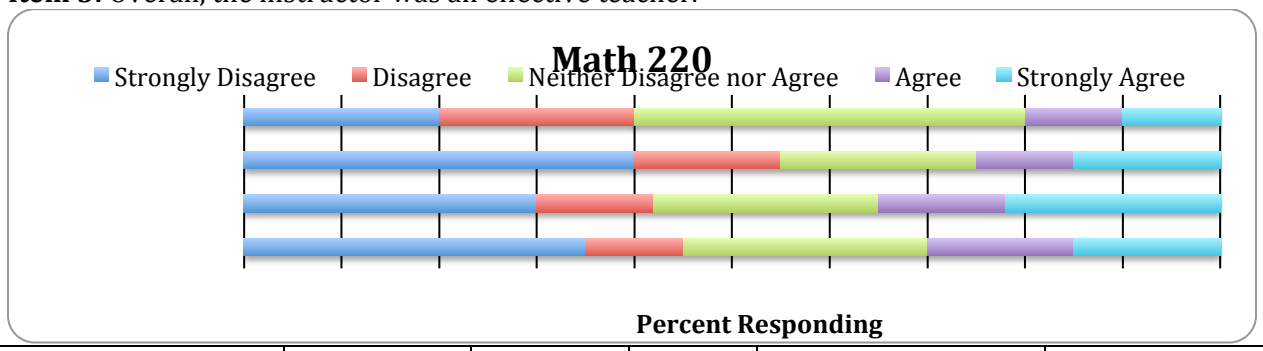
Math 220: Elementary Statistics
Instructor: John Brown
Semesters: Fall 2011-Spring 2012

Item 1: The instructor respects students.



	No. Responses	Average	Std. Dev.	%Agree or Strongly Agree	%Disagree or Strongly Disagree
Fall 2012 (Sec 1; n=50)	50	2.65	1.465	30	45
Fall 2012 (Sec 2; n=30)	30	2.85	1.526	35	42
Spring 2012 (n=30)	30	2.45	1.472	25	55
Fall 2011 (n=34)	34	2.7	1.19	20	40

Item 3: Overall, the instructor was an effective teacher.



	No. Responses	Average	Std. Dev.	%Agree or Strongly Agree	%Disagree or Strongly Disagree
Fall 2012 (Sec 1; n=50)	50	2.65	1.465	30	45
Fall 2012 (Sec 2; n=30)	30	2.85	1.526	35	42
Spring 2012 (n=30)	30	2.45	1.472	25	55
Fall 2011 (n=34)	34	2.7	1.19	20	40

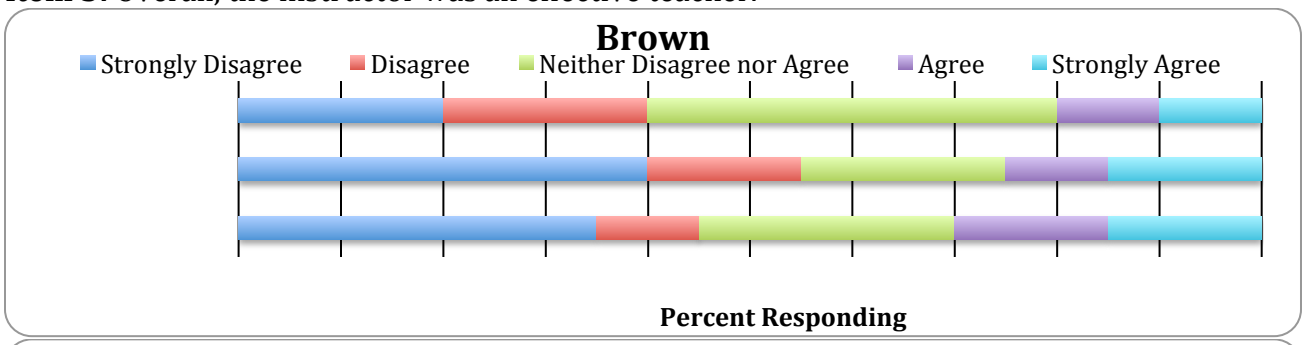
SAMPLE DEPARTMENTAL AGGREGATE REPORT- VERSION A

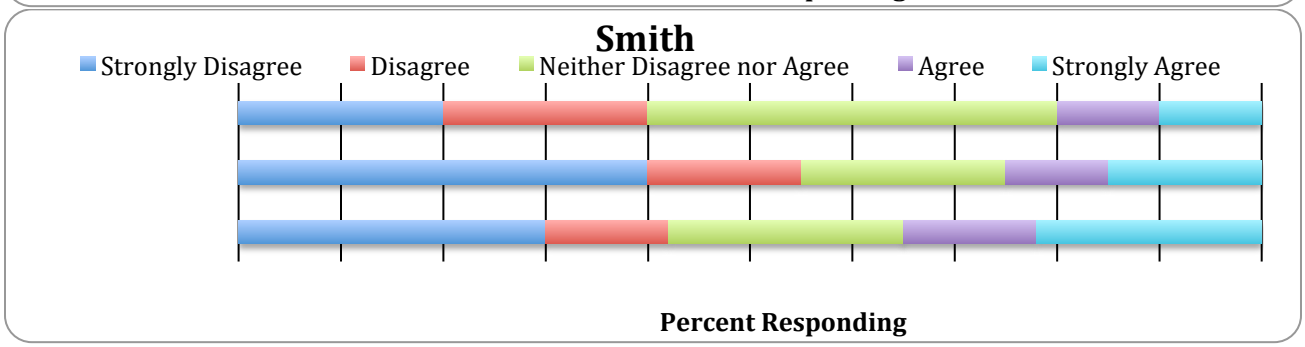
The goal of such a report is to compare different instructors of the same course in a single semester and/or summarize the SET for each individual instructor. Items selected for comparison should be equally appropriate to each instructor's teaching style and each course being compared and should be identically worded with the same Likert scale used in each.

SET Summary Section

Semester: Spring 2012
Instructors: Shoney, MacGuire, Hayworth, Jones, Smith and Brown

Item 3: Overall, the instructor was an effective teacher.





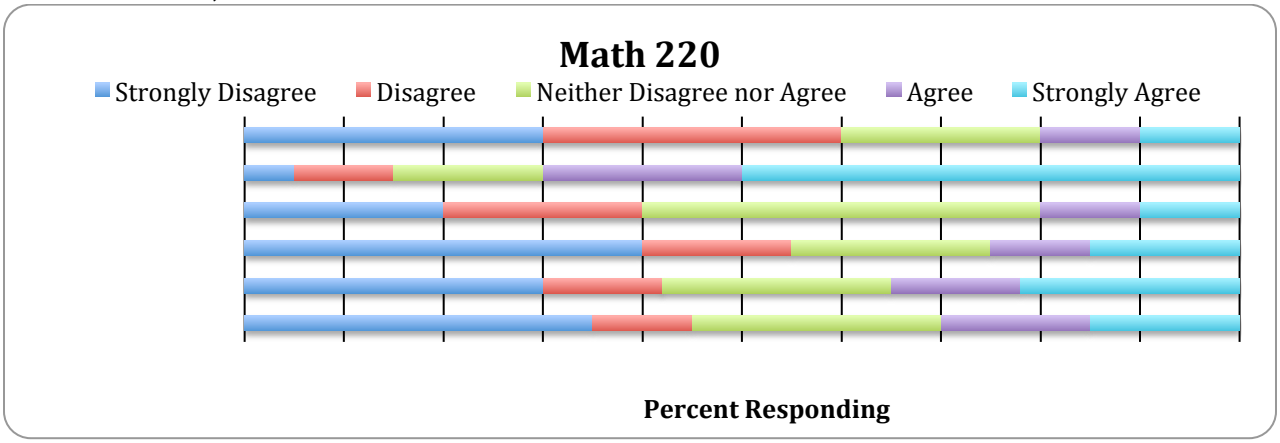
Brown	No. Responses	Average	Std. Dev.	%Agree or Strongly Agree	%Disagree or Strongly Disagree
Math 220-001	34	2.65	1.47	30	45
Math 220-002	34	2.45	1.47	25	55
Math 280	22	2.7	1.19	20	40

Smith	No. Responses	Average	Std. Dev.	%Agree or Strongly Agree	%Disagree or Strongly Disagree
Math 220-011	30	2.65	1.47	30	45
Math 235-002	20	2.45	1.47	25	55
Math 429	4	2.7	1.19	20	40

Instructor Comparison Section

Math 220: Elementary Statistics
Semester: Spring 2012
Instructors: Shoney, MacGuire, Hayworth, Jones, Smith and Brown

Item 3: Overall, the instructor was an effective teacher.



Math 220	No. Responses	Average	Std. Dev.	%Agree or Strongly Agree	%Disagree or Strongly Disagree
Brown	100	2.65	1.47	30	45
Smith	100	2.85	1.53	35	42
Jones	100	2.45	1.47	25	55
Hayworth	100	2.70	1.19	20	40
MacGuire	100	4.00	1.23	70	15
Shoney	100	2.40	1.29	20	60

SAMPLE DEPARTMENTAL AGGREGATE REPORT – VERSION B

The goal of such a report is to summarize the SET for each individual instructor. Items selected for comparison should be equally appropriate to each instructor's teaching style and each course being compared and should be identically worded with the same Likert scale used in each.

Semester: Spring 2012
Instructor: Smith

Math 220 - 001	**No. Responses**	**Average**	**Std. Dev.**	**%Strongly Disagree**	**%Disagree**	**%Neither Disagree nor Agree**	**%Agree**	**%Strongly Agree**
Item 1: Clear expectations	34	3.65	1.23	20	15	30	20	15
Item 2: Respects students	33	3.6	1.23	20	15	30	20	15
Item 3: Overall effective teacher	34	2.65	1.07	20	15	30	20	15

Math 220 - 002	**No. Responses**	**Average**	**Std. Dev.**	**%Strongly Disagree**	**%Disagree**	**%Neither Disagree nor Agree**	**%Agree**	**%Strongly Agree**
Item 1: Clear expectations	34	3.65	1.23	20	15	30	20	15
Item 2: Respects students	33	3.6	1.23	20	15	30	20	15
Item 3: Overall effective teacher	34	2.65	1.07	20	15	30	20	15

Math 28	**No. Responses**	**Average**	**Std. Dev.**	**%Strongly Disagree**	**%Disagree**	**%Neither Disagree nor Agree**	**%Agree**	**%Strongly Agree**
Item 1: Clear expectations	34	3.65	1.23	20	15	30	20	15
Item 2: Respects students	33	3.6	1.23	20	15	30	20	15
Item 3: Overall effective teacher	34	2.65	1.07	20	15	30	20	15

Semester: Spring 2012
Instructor: Brown

Math 220 - 001	**No. Responses**	**Average**	**Std. Dev.**	**%Strongly Disagree**	**%Disagree**	**%Neither Disagree nor Agree**	**%Agree**	**%Strongly Agree**
Item 1: Clear expectations	34	3.65	1.23	20	15	30	20	15
Item 2: Respects students	33	3.6	1.23	20	15	30	20	15
Item 3: Overall effective teacher	34	2.65	1.07	20	15	30	20	15

Math 220 - 002	**No. Responses**	**Average**	**Std. Dev.**	**%Strongly Disagree**	**%Disagree**	**%Neither Disagree nor Agree**	**%Agree**	**%Strongly Agree**
Item 1: Clear expectations	34	3.65	1.23	20	15	30	20	15
Item 2: Respects students	33	3.6	1.23	20	15	30	20	15
Item 3: Overall effective teacher	34	2.65	1.07	20	15	30	20	15

Math 28	**No. Responses**	**Average**	**Std. Dev.**	**%Strongly Disagree**	**%Disagree**	**%Neither Disagree nor Agree**	**%Agree**	**%Strongly Agree**
Item 1: Clear expectations	34	3.65	1.23	20	15	30	20	15
Item 2: Respects students	33	3.6	1.23	20	15	30	20	15
Item 3: Overall effective teacher	34	2.65	1.07	20	15	30	20	15

VI. Implementation Subcommittee Report

Members

- **Christie Liu**
- **Catherine Crummett**
- **Andreas Broscheid**
- **Kristin Sowden**

1. Timeline

1.1 Goal

The goal is to have all academic units utilizing the Blue system for student course evaluations by the fall 2015 semester.

1.2 Spring 2013 & Fall 2014 Pilots

Approximately one-half of the academic units currently using online evaluations will participate in the spring 2013 pilot. These academic units currently use either Blackboard or Qualtrics for evaluations. The remaining half will participate in the fall 2014 pilot along with departments using paper evaluations who volunteer for the fall 2014 pilot. The following departments are confirmed for participation in one of the two semesters include:

- Biology
- College of Education – all units
- History
- Honors Program
- Math
- Military Science
- Nursing – NSG 653
- Physics
- Psychology
- School of Engineering
- Social Work - TBD
- WRTC

The table below outlines meetings with the participating academic units and the system process.

Date	Meeting/Function	Purpose	Responsibility
Sep 2012	Meet with academic units	Informational meeting for faculty and AUH	Steering committee and Provost
Nov 2012	Spring 2013 pilot group unit representative(s) meeting	Share unit experiences with online evaluation; learn about the Blue system; learn about the JMU online evaluation process.	Implementation subcommittee
Nov 2012	Online training available		CIT
Feb 1	Academic unit questions sent to the System Administrator		System Administrator
Mar 1	Blue opens to faculty for faculty question entry	Faculty enter individual questions or select from a question bank.	Faculty
Mar 29	Blue closes to faculty	Questions may no longer be entered by faculty. If necessary, the Blue system administrator may be able to add questions.	
Arp 17	Evaluations open		System Administrator
Apr 17 – May 3	Email reminders are sent to students	Emails will be sent every 3 days.	System Administrator
May 3 midnight	Evaluations close		System Administrator
May 9	Faculty reports are available in *MyMadison*	Reports are available 48 hours after grades are submitted	System Administrator
TBA	Debriefing/Feedback Meeting	Meet with faculty and AUH to obtain feedback from the pilot.	Standing committee

1.3 Preparation for future pilots (Fall 2014, Spring 2015 and Fall 2015)

Academic units will be asked to volunteer to participate in the fall 2013, spring 2014 and fall 2015 pilots following the process outline of the table above. All remaining academic units will participate by the spring 2015 semester.

2 Feedback mechanisms

2.1. Meetings with department faculty, including chairs, at the beginning of the semester of first use of the eXplorance Blue system

Purposes:

- Discuss advantages of the new system and faculty concerns.
- Discuss best practices based on research.
- Work with faculty and department head to make sure evaluation process meets faculty and department needs.

2.2. Faculty, chair, and student online surveys conducted at the beginning of the semester after evaluation. The three stakeholder groups will receive separate surveys, though questions may overlap. The surveys will be designed by the working group or a subcommittee and administered through Qualtrics.

Purposes:

- Obtain information on potential implementation problems and identify areas for improvement in the next implementation cycle.
- Determine to what extent the process has met the needs of the three stakeholder groups and identify improvement for the next implementation cycle.

2.3. Feedback meetings with department faculty and chair during weeks 2-5 of semester after evaluation

Purposes:

- Obtain feedback and suggestions from the AUH and faculty.
- Work with faculty and chairs to identify improvements for next implementation cycle.

2.4. Student focus groups in the semester after the evaluation.

Participants are identified through the student survey. The student survey includes a question that asks whether students were willing to participate in a focus group. Participants are then selected to be roughly representative of the student body.

Purposes:

- Identify strengths and shortcomings of eXplorance/Blue from students' perspective.
- Identify improvements for the next implementation cycle.
- Obtain student perspective on new evaluation technology.

2.5. Regular meetings of the IT Group

Purposes:

- Discuss changes to use of eXplorance/Blue identified in surveys, department meetings, and focus groups.
- Minutes from each meeting will serve to document implementation issues.

2.6. Regular meetings of the standing committee (see point 4)

3 Communication Strategy

3.1 Faculty:

- Website (Spring 2013)
  - General information
  - FAQs
  - Training tutorial links
  - Document on questionnaire best practices
  - Document on best practices regarding SET use, including data analysis and suggestions for how to interpret the data
  - Pilot success stories
  - Links to academic research regarding online evaluation systems
- Email from Provost – thanking academic units for participating in the current pilot
- Meetings - meet with participating departments (see 2.1 and 2.2)
- Faculty Senate - updated regularly or during the first meeting of each semester

3.2 Students (Spring 2013):

- Inform SGA
- Website – General information, FAQs
- Email from Registrar's Office – announcement of the new system and the system process, the transition period and the variety of evaluation systems they may be asked to used until the transition is complete (Qualtrics, Blackboard, ScanTron, Blue, etc.)
- Breeze article – announcement of the new system and the system process
- Faculty announcement – faculty who discuss the importance of the course evaluation and how the information is used to improve the course proves to be the most influential factor in increasing and maintaining response rates.
- Student Survey/Focus group - information regarding strengths/weaknesses of the Blue system based; suggest improvements to increase response rates

4 Standing committee

4.1 Responsibilities. The standing committee should have the following responsibilities:

- meet with departments for information and feedback
- analyze feedback surveys

- identify (non-technical) implementation problems and make recommendations for solutions
- provide input to IT group
- maintain and circulate information to faculty, departments, and students

4.2 Representation. The standing committee should include members representing

- CFI
- CIT
- faculty including assistant, associate, and full professors
- non-tenure-track faculty
- dept chairs
- CARS
- members of several colleges
- University Studies
- Student representatives – SGA

5. Roles

5.1 Academic Unit Head

- Provide current evaluation
- Submit modifications to evaluations as necessary
- Provide input on academic unit level questions
- Serve as/appoint a liaison to work with the standing committee and faculty
- Communicate with faculty

5.2 Administrative Assistants

At this point in time, modifications to the evaluation will be made by the system administrator. Administrative Assistants have no responsibility with regard to the Blue system.

5.3 Faculty

- Input formative questions within the designated time period
- Encourage student response by stressing the importance of student feedback and how the feedback is used to improve or modify the course
- Review reports
- Send summative date to other units as necessary (GenEd, supervisor, etc.)

5.4 System Administrator

- Upload/modify evaluations
- Date administration – set dates for opening and closing the system

- Send email reminders after the first two automated emails
- Send SA data to eXplorance
- Responsible for the semester system processes overall

VII. Evaluating Teaching Subcommittee Report

Members

- Morgan Benton
- Mark Parker
- Adebayo Ogundipe
- Maggie Kyger

Committee Members:

- Morgan Benton, Assistant Professor, ISAT
- Maggie Kyger, Assistant Dean, College of Education
- Adebayo Ogundipe, Assistant Professor, Engineering
- Mark Parker, Department Head, English

Goal of the Evaluating Teaching Subcommittee

The goal of the Course Evaluation Subcommittee is to take a broad and holistic look at the practice of teaching and to make recommendations to the SET Task Force with regard to the SET instrument being developed that will place it properly within a ***context*** of evaluation of teaching as a whole. The Subcommittee will explore and examine a broad range of methods for evaluating teaching and provide guidance to departments in the development and/or evolution of their strategies for evaluating teaching. Specifically, the Subcommittee will address:

- What are the dimensions of high quality teaching?
- What types of data might be gathered by departments and professors to shed light on the quality of teaching being performed by professors?
- What role should the Student Evaluation of Teaching (SET) instrument play in the context of a broader evaluation of teaching strategy?

To a lesser extent, the Subcommittee will also comment on how any collected data should be handled and interpreted for both summative and formative purposes.

Organization of this Report

This report begins with a short "beliefs and values" statement followed by a statement of the Subcommittee's view of what constitutes high quality teaching. Following that, the report is broken up into two sub-sections concerning the dimensions of high quality teaching and the types of data that can be gathered to make an informed analysis of the quality of actual teaching. The report concludes with a table suggesting elements that together would comprise a teaching portfolio that instructors may use for both formative and summative evaluation of their own teaching.

Beliefs and Values of the Subcommittee

The Evaluation of Teaching Subcommittee believes:

- Teaching is a complex, multi-dimensional practice which cannot be evaluated effectively by any single instrument, whether that be student evaluation, peer observation, personal reflection or otherwise
- That being said, students are critical stakeholders in the process of teaching and learning and have valuable perspectives that can be used in the development of an instructor's teaching practice
- Increased student learning, another complex and difficult to assess phenomenon, is of the highest priority and the primary goal of efforts to evaluate teaching
- All instructors have room for improvement—as instructors, we are never "done" when it comes to developing our craft
- The overwhelming majority of instructors at JMU are committed and conscientious practitioners of their craft currently at different stages in their personal growth and development
- Consequently, the instances in which quality of teaching falls so low as to warrant removal/suspension of an instructor from teaching roles are ***extremely rare*** though not nonexistent
- The quality of student experience and of instructors' professional lives could be strengthened significantly by a thoughtful, holistic approach to the evaluation of teaching

What Constitutes High Quality Teaching?

Teaching is a complex, multi-dimensional, professional practice that defies simple, one-size-fits-all definitions. While a deep knowledge in one's chosen field is an important dimension of quality teaching, it is only one component—quality teaching also comprises understanding of how humans learn, diverse pedagogical strategies that can be used to engage students, an appropriate attitudinal disposition towards students and teaching, as well as professional habits and behaviors associated with rich learning environments. The knowledge and skills one needs to be a high quality instructor are not always a part of the deeper training that instructors receive in their fields (e.g. as a part of graduate school) and must be developed "on the job" during one's career.

Dimensions of High Quality Teaching

The following sub-section describes the dimensions of high quality teaching. The categories listed here are not necessarily comprehensive nor mutually exclusive, but represent the Subcommittee's first attempt at exploring this concept.

Knowledge of Content (KoC)

A high quality instructor should be an expert in the content being taught. The standards for what constitutes an "expert" will naturally vary across fields and disciplines. Content may include skills, objective knowledge, modes of inquiry or problem solving, and attitudes of practitioners in a given domain.

Knowledge of Pedagogy (KoP)

Content may be taught in a variety of ways. High quality instructors are aware of the various pedagogical tools and teaching practices that have and can be used to create a high quality learning experience for students. These practices will vary widely given the type of content, size of classes, type of students, and other factors related to the specific context of the teaching situation. Underlying many or most of these practices is an understanding about how humans learn in general, and more particularly how they best learn a particular body of content.

Knowledge of Assessment (Assessment)

An instructor's ability to provide high quality feedback to students on their learning processes is predicated upon their ability to gather high quality information about what students are learning as they engage with content. High quality instructors, therefore, will be aware of and employ a range of assessment tools that allow them to understand where students stand, and in turn convey that status to them.

Student Learning (Learning)

In the presence of high quality instructions, learning gains by students are higher. What students learn and how this is assessed are questions not dealt with by this particular dimension of teaching.

Course Preparation (Prep)

High quality instructors put significant energy into preparing for teaching.

Student Expectations (Expectations)

High quality instructors not only have high expectations of their students, but also make efforts to make course learning objectives relevant to students. This may involve changing the objectives, or it may involve changing the perspectives of students so that they come to see course learning objectives as relevant.

Respect for Students (Respect)

This category actually captures a collection of attitudes and behaviors that high quality instructors exhibit towards students including: trust, respect, decency, time commitment, consistency, flexibility, and not blaming students for having arrived in a class inadequately prepared by past experiences/coursework.

Post-class Achievements (Achievement)
As a result of high-quality instruction, students will go on to achieve notable things in the future. This is not assessable within the confines of the semester in which a course is taught but entails an effort to keep track of what students achieve once they have left the classroom and are asked to apply their knowledge and skills in their lives.

Self-Reflection (Reflection)
High quality instructors reflect on their own performance and are continually engaged in a process that will allow them to grow and adapt as instructors over the lifetime of their careers.

Modeling Life-long Learning (Modeling)
High quality instructors are models for students of what we mean by "life-long learning." They convey to students that the goal of learning is not solely for the purpose of receiving a passing grade or degree credentials, but that it is the path towards becoming an "enlightened citizen who leads a meaningful and productive life."

Commitment to Academic Community (Commitment)
High quality instructors exhibit a commitment to furthering the scholarship in their field.

Types of Data That Can Inform the Teaching Evaluation
Student perceptions of teaching can not and should not be expected to provide information about all of the aspects of high quality teaching described above. The Subcommittee rather envisions that instructors and departments will collect a range of data that collectively can inform the teaching evaluation. Again, this is not intended to be a comprehensive list, but rather a suggestion of some of the more obvious approaches to gathering data on teaching.

Student Perceptions of Teaching (Students)
These are instruments delivered either during or at the end of teaching which gather students' perceptions of the teaching and learning process. The SET instrument that this Task Force is working to create would fall into this category.

Peer Observation and/or Peer-Facilitated Reflection (Peer)
This category comprises situations in which one or more peers actually sit in and observe an instructor/colleague at work, or methods such as the Teaching Analysis Poll (TAP) in which a colleague will facilitate a discussion with students designed to gather feedback about the teaching and learning process.

Instructor Narrative/Self-Reflection (Self)
Instructors can and should reflect upon their practice as teachers. This reflection can be more or less formal, and may manifest itself in more or less concrete goals for intentional self-development. These goals will change over time as the instructor matures as a teacher.

Objective Measures of Learning (Objective)
Many fields have standardized tests, concept inventories, or other assessments which can provide an externally validated indication of the levels of student learning in a particular subject area.

Scholarly Work (Scholarship)
Instructors can demonstrate knowledge of subject matter through their publications, creations, consulting or outreach work, or in other ways that would typically be classified as "scholarly work" for the purposes of promotion and tenure.

Measures of Students' Metacognitive Growth (Metacognition)
Metacognition refers to all of the ways that students may grow in a course other than with respect to the core content covered in the course. This could be changes in attitudes, beliefs, behaviors, and motivation. It could also comprise measures of the "21st Century Skills": creativity, collaboration, communication, critical thinking.

Portfolios of Student Work (Portfolios)
Another way to demonstrate level and quality of student work over time is to collect representative samples of student work that can be assembled into portfolios that highlight the level of quality students are capable of after having studied under an instructor.

Syllabi and Other Course Materials (Syllabi)
Instructors work hard to produce high quality syllabi and other supporting documents for a course such as engaging assignments, tutorials, course readers, and labs. These materials can provide evidence of meticulous course preparation.

Summary
The table on the following page shows the relationship between the dimensions of high quality teaching and the sources of data that would inform the evaluation of teaching. It is intended to put the SET instrument in context and also to provide guidance to instructors and departments engaged in the process of developing an evaluation strategy for teaching.

Table 1: Ways that various types of data can be used to evaluate the quality of teaching

Type of Data	Dimensions of High Quality Teaching										
	KoC	KoP	Assessment	Learning	Prep	Expectations	Respect	Achievement	Reflection	Modeling	Commitment
Students					✔	✔	✔			✔	
Peer	✔	✔					✔				
Self	✔	✔	✔		✔	✔	✔		✔	✔	✔
Objective				✔				✔			
Scholarship	✔									✔	✔
Metacognition				✔							
Portfolios		✔	✔	✔		✔		✔			✔
Syllabi	✔	✔	✔		✔	✔	✔		✔		

VIII. Default Evaluation Items

Recommended Default Student Evaluation

1. I learned a significant amount in this course.

2. I attended class regularly.

3. On average, I spent _____ hours per week working on this course outside of class.
0-2 2-4 4-6 6-8 8+

4. The instructor stimulated my interest in the subject.

5. Overall, the instructor was an effective teacher.

6. The instructor was available during scheduled office hours.

7. The instructor seemed enthusiastic about teaching the course.

8. The instructor conveyed clear expectations for all graded work.

9. The instructor provided timely feedback on graded work.

10. The instructor adhered to the syllabus.

11. The instructor organized the course effectively.

12. The instructor was fair and impartial in assessing student work.

13. What characteristics of the course were **most** valuable to your learning?

14. What characteristics of the course were **least** valuable to your learning?

15. Feel free to write any other comments about the instructor or the course:

IX. Evaluation Item Pool

I. Questions about the Course

The questions in this section assume a Likert scale with five responses: Strongly Disagree, Disagree, Neither Agree nor Disagree, Agree, Strongly Agree. Each question should also have an "opt out" mechanism.

1. The course was intellectually challenging.
2. The course changed the way I think.
3. The course increased my interest in the topic.
4. The course gave me an appreciation of the subject matter.
5. The course enhanced my learning.
6. The course helped me learn fundamental aspects of the discipline.
7. The course helped me gain factual knowledge.
8. The course improved my communication skills.
9. The course improved my writing skills.
10. The course improved my oral communication skills.
11. The course improved my laboratory skills.
12. The course improved my analysis and design skills.
13. The course improved my problem solving skills.
14. The course improved my teamwork skills.
15. The course improved my research skills.
16. The course helped me learn to evaluate ideas, arguments, and points of view critically.
17. The course helped me develop a clearer understanding of my values.
18. The course helped me develop clearer understanding of ethical issues in the profession.
19. The course helped me develop clearer understanding of citizenship and social responsibility.
20. The course evaluation criteria were clearly stated in the syllabus.
21. The course covered a reasonable amount of material.
22. The course assignments enhanced my learning.
23. The instructional material was valuable.
24. The textbook was valuable.
25. The course tests assessed my learning accurately.
26. The course was organized effectively.
27. I learned a significant amount in this course.

II. Questions about the Instructor

The questions in this section assume a Likert scale with five responses: Strongly Disagree, Disagree, Neither Agree nor Disagree, Agree, Strongly Agree. Each question should also have an "opt out" mechanism.

28. Overall, the instructor was an effective teacher.
29. The instructor encouraged me to reflect on my own performance.
30. The instructor had a genuine interest in individual students.
31. The instructor demonstrated respect for students.
32. The instructor demonstrated a thorough grasp of the course material.
33. The instructor was well prepared for classes.
34. The instructor helped students answer their own questions.
35. The instructor answered questions clearly.
36. The instructor seemed enthusiastic about teaching the course.
37. The instructor demonstrated the importance of the subject matter.
38. The instructor made it clear how each topic fit into the course.
39. The instructor explained material in an interesting manner.
40. The instructor stimulated my interest in the subject.
41. The instructor explained course material clearly.
42. The instructor made interdisciplinary connections.
43. The instructor brought outside speakers to the class.
44. The instructor held class as scheduled.
45. The instructor gave assignments related to important material in the course.
46. The instructor provided timely feedback on graded work.
47. The instructor provided frequent assignments.
48. The instructor provided appropriate assignments.
49. The instructor was available during scheduled office hours.
50. The instructor was generous with his or her time outside of class.
51. The instructor explained the grading system used in the class clearly and completely.
52. The instructor behaved professionally.
53. The instructor provided work outside of class that was valuable for learning the course material.
54. The instructor gave examinations that covered the most important course material.
55. The instructor used evaluation methods appropriate to the course.
56. The instructor encouraged students to express their own ideas and to ask questions.
57. The instructor stressed important points.

58. The instructor defined new or unfamiliar terms.
59. The instructor explained difficult concepts several times or in several ways.
60. The instructor pointed out practical applications of the material.
61. The instructor was an engaging lecturer.
62. The instructor challenged me to think critically.
63. The instructor used a variety of visual and other aids.
64. The instructor used a variety of activities during class.
65. The instructor did not have distracting mannerisms.
66. The instructor encouraged class discussions.
67. The instructor used class time efficiently.
68. The instructor led class discussion effectively.
69. The instructor offered constructive criticism.
70. The instructor adhered to the syllabus.
71. The instructor organized the course effectively.
72. The instructor explained the organization of the course.
73. The instructor delivered well-organized lectures.
74. The instructor advised students how best to prepare for and complete assignments.
75. The instructor advised students how best to prepare for examinations.
76. The instructor conveyed clear expectations for all graded work.
77. The instructor expressed clear expectations for my learning and performance.
78. The instructor had high standards for student achievement.
79. The instructor had reasonable standards for student achievement.
80. The instructor reminded students of test dates or assignment deadlines.
81. The instructor was fair and impartial in assessing student work.

III. Questions about the Student

The first seven questions in this section assume a Likert scale with five responses: Strongly Disagree, Disagree, Neither Agree nor Disagree, Agree, Strongly Agree. The remaining questions have individual responses. In most cases, question should also have an "opt out" mechanism.

82. I attended class regularly.
83. I came to class prepared.
84. I participated in discussions.
85. I put a great deal of effort into advancing my learning in this course.
86. I learned more than I expected to learn in this course.
87. I worked up to my potential in this course.

88. I believe that what I was asked to learn in this course is important.
89. On average, I spent _____ hours per week working on this course outside of class.
    0-2 2-4 4-6 6-8 8+
90. I expect to receive a grade of _____ in this course.
    F D- to D+ C- to C+ B- to B+ A- to A
91. My gender is _____ (Male of Female).
92. My current class status is _____ (Freshman, Sophomore, Junior, Senior).
93. This course:
    is required for my major or minor
    is an elective in my major or minor
    is an elective outside my major
    meets a General Education requirement
    meets an IDLS requirement.
94. I have been absent from this class about _____ times.
95. I have met with the instructor outside of class about _____ times.
96. My GPA is
    0.0 to 1.5 1.51 to 2.5 2.51 to 3.5 3.51 to 4.

IV. Course Comparison Questions

The questions in this section assume a Likert scale with five responses: Much Less, Less, About the Same, More, Much More. All question should also have an "opt out" mechanism.

97. Compared to other general education courses, the quality of this course was
98. Compared to other general education courses, the amount I learned in this course was
99. Compared to other general education courses, the workload in this course was
100. Compared to other general education courses, the amount of reading in this course was
101. Compared to other general education courses, the amount of writing in this course was
102. Compared to other general education courses, the amount of teamwork in this course was
103. Compared to other general education courses, the amount of laboratory work in this course was
104. Compared with other courses in general education, the amount of homework in this course was
105. Compared to other general education courses, the amount of project work in this course was
106. Compared to other major or minor courses, this course was
107. Compared to other major or minor courses, the amount I learned in this course was
108. Compared to other major or minor courses, the workload in this course was

109. Compared to other major or minor courses, the amount of reading in this course was
110. Compared to other major or minor courses, the amount of writing in this course was
111. Compared to other major or minor courses, the amount of teamwork in this course was
112. Compared to other major or minor courses, the amount of laboratory work in this course was
113. Compared to other major or minor courses, the amount of homework in this course was
114. Compared to other major or minor courses, the amount of project work in this course was

V. Instructor Comparison Questions

The questions in this section assume a Likert scale with five responses: Much Worse, Worse, About the Same, Better, Much Better. All question should also have an "opt out" mechanism.

115. Compared with other instructors, this instructor was …
116. Compared with other instructors, this instructor's lecture style was …
117. Compared with other instructors, this instructor' ability to answer questions was …
118. Compared with other instructors, this instructor's feedback on assignments was …
119. Compared with other instructors, this instructor' availability was …

VI. Open-Ended Questions

120. What characteristics of the course were most valuable to your learning?
121. What characteristics of the course were least valuable to your learning?
122. What characteristics of the instructor were most valuable to your learning?
123. What characteristics of the instructor were least valuable to your learning?
124. How could this course be improved?
125. How could course materials be improved?
126. How could assignments be improved?
127. How could course activities be improved?
128. How could assessment or grading be improved?
129. How could projects be improved?
130. How could laboratories be improved?
131. How could lectures be improved?
132. List the two or three assignments or topics from which you learned the most.
133. List the two or three assignments or topics from which you learned the least.

134. Why did you take this class?

135. Feel free to write any other comments about the instructor or the course.

VI. Questions About the Pilot Survey

136. What questions in this survey were hard to understand?

137. What questions in this survey should not have been asked?

138. What questions are missing from this survey?

Guidelines for Constructing your Own Questions

1. Keep your evaluation to a reasonable length. Excessive evaluation length (more than 15-20) items may reduce the validity of the results.
2. Write simple, positive, declarative sentences in past tense using the active voice.
3. Use simple vocabulary and sentence structure.
4. Address only one issue per item. ("The instructor explains concepts clearly and concisely" would be problematic, as it addresses two distinct issues.)
5. Don't write items that conflate course content with teaching; items should address one or the other.
6. Don't ask questions that students are not competent to answer (for example, "The instructor is an expert in her field").
7. Leave space for comments for as many questions as possible.
8. Whenever possible, use a consistent point scale and a consistent scale direction throughout the survey.
9. Avoid questions that imply an endorsement of one particular teaching style or strategy (for example, "The instructor included plenty of group work" or "The instructor moved around the room as he or she spoke").