

University of Rochester Course Evaluation Project

Ronald D. Rogge

Associate Professor

Ista Zahn

Doctoral Candidate

Department of Clinical & Social Sciences in Psychology

## **Project Impetus**

The current online administration of course evaluation surveys at the University of Rochester has led to notably lower response rates, dropping from the 80-90% rates typically observed from in-class administration to 30-40%, despite efforts by the administration to provide incentives to promote higher participation. Such a marked drop in participation could effectively render the information obtained by those online course evaluations less reliable and/or excessively biased. To address this issue, we conducted a study to directly compare the quality of information obtained from online vs. in-class administration of course evaluations. However, to thoroughly examine that more focused goal it was advantageous to examine a broader set of 5 inter-related goals.

## **Project Aims**

**Goal 1: Examining the Current Questions.** The project examined the questions with numeric responses currently being used in the University of Rochester's course evaluation to determine how much unique information they offer instructors.

**Goal 2: Evaluating Bias.** The project sought to determine the degree to which known biases influence responses on each of the course evaluation questions being used at the University of Rochester.

**Goal 3: Developing a New Tool.** The project sought to develop a new evaluation tool that could offer instructors more useful and diverse feedback with markedly lower levels of bias.

**Goal 4: Contrasting Online to In-Class Administration.** The project sought to determine if the lower response rates seen with online administration might lead to unreliable or excessively biased data.

**Goal 5: Determining Number of Responses Needed.** The project sought to determine the minimum number of student responses necessary to obtain reasonably accurate and reliable estimates of course ratings for an individual course.

## **Project Method**

**Overall Design.** The project collected course evaluation data from 1,519 students from 48 courses across 20 departments in the Spring 2010 semester. Students completed the course evaluation questions currently in use at the University of Rochester as well as a set of 80 additional items that included: 1) a diverse pool of course evaluation items currently in use at other universities around the globe, and 2) items assessing potential sources of bias in course ratings.

**Selecting Courses.** The courses were specifically selected to represent the diversity of courses offered at the University of Rochester. Courses were selected in pairs from each department (matching on course size and course level) and then courses within each pair were randomly assigned to administer course evaluations in-class or via the current online system. This helped to ensure that online vs. in-class

comparisons would be made on comparable sets of courses. The project was approved by the Research Subjects Review Board at the University of Rochester and all students and instructors involved were informed of their rights prior to consenting to the study. We achieved a response rate of 67% using in-class administration and a response rate of 40% using online administration, mirroring the difference in response rates seen at the university level since instituting the online administration system.

**Biases Influencing Responses.** After a thorough review of the literature in this area and some preliminary analyses, we focused the project onto the following set of biases: student gender, student personality traits (extraversion, agreeableness, conscientiousness, openness, neuroticism), expected grade, GPA, prior interest in course, perceived course difficulty, and perceived instructor attractiveness. These were assessed with standard one or two item measures.

## Project Results

**Goal 1: Examining the Current Questions.** The project used exploratory factor analyses (EFA)<sup>1</sup> on the questions (with numeric responses) currently being used in the University of Rochester's course evaluation to determine how many dimensions of information (factors) they represent. The EFA analyses on the items currently in use at the University of Rochester suggested that they were essentially measuring four distinct constructs:

- Overall Quality
  - A set of 5 strongly correlated items assessing global ratings of course quality
    - What overall rating would you give this course?
    - What overall rating would you give this instructor?
    - How effective was the instructor's teaching in this course?
    - Rate the increase of your knowledge or skills from this course.
    - I have a stronger interest in this subject because of the instructor.
- Student Effort
  - 2 strongly correlated items assessing students' reports of the effort they expended
    - Rate the level of your involvement in the activities of this course (for example: attendance, participation, completing assignments).
    - What overall rating would you give yourself as a student in this course?
- Quality of Readings
  - 1 item assessing the quality of course readings
    - The readings were important in my learning in this course.
- Quality of Syllabus
  - 1 item assessing the quality of course syllabus
    - How well did the syllabus describe the course content?

Based on these EFA results, we created composite scores representing overall quality and student effort so that the remaining analyses could examine the factors identified. In addition, although instructors are given

---

<sup>1</sup> This technique examines sets of items to help researchers identify subsets of items that seem to be measuring the same underlying construct (the same factor). As a result, EFA enables researchers to take a larger set of items and simplify them down into a handful of dimensions represented by those items. After identifying the subsets of items that make up each dimension, researchers then examine the item content within each of those dimensions to come up with descriptive labels to represent what each set of items seem to measuring.

feedback on student responses to all questions on the current evaluation form, there are two questions (What overall rating would you give this course? What overall rating would you give this instructor?) that are being used to represent course quality in faculty activity reports. Given the emphasis placed on these two single items, we examined the quality of information offered by these individual items in the remaining analyses as well.

**Goal 2: Evaluating Bias.** To determine the degree to which student responses were unduly influenced by factors other than course quality, we ran analyses<sup>2</sup> in which we allowed a comprehensive set of biases to predict student responses on each of the items of the current evaluation form. After a thorough review of the literature and some preliminary analyses, we focused the project onto the following biases:

- Student gender
- Student personality traits
  - Extraversion                      Agreeableness                      Emotional Stability
  - Conscientiousness              Openness
- Student's expected grade
- Student GPA
- Student's prior interest in course
- Student perceptions of course difficulty
- Student perceptions of instructor attractiveness

We also examined a number of additional sources of bias. However, after controlling for the biases listed above, these remaining biases failed to demonstrate any influence on student responses to course evaluation items, and so, in the interest of parsimony, they were dropped from the final analyses.

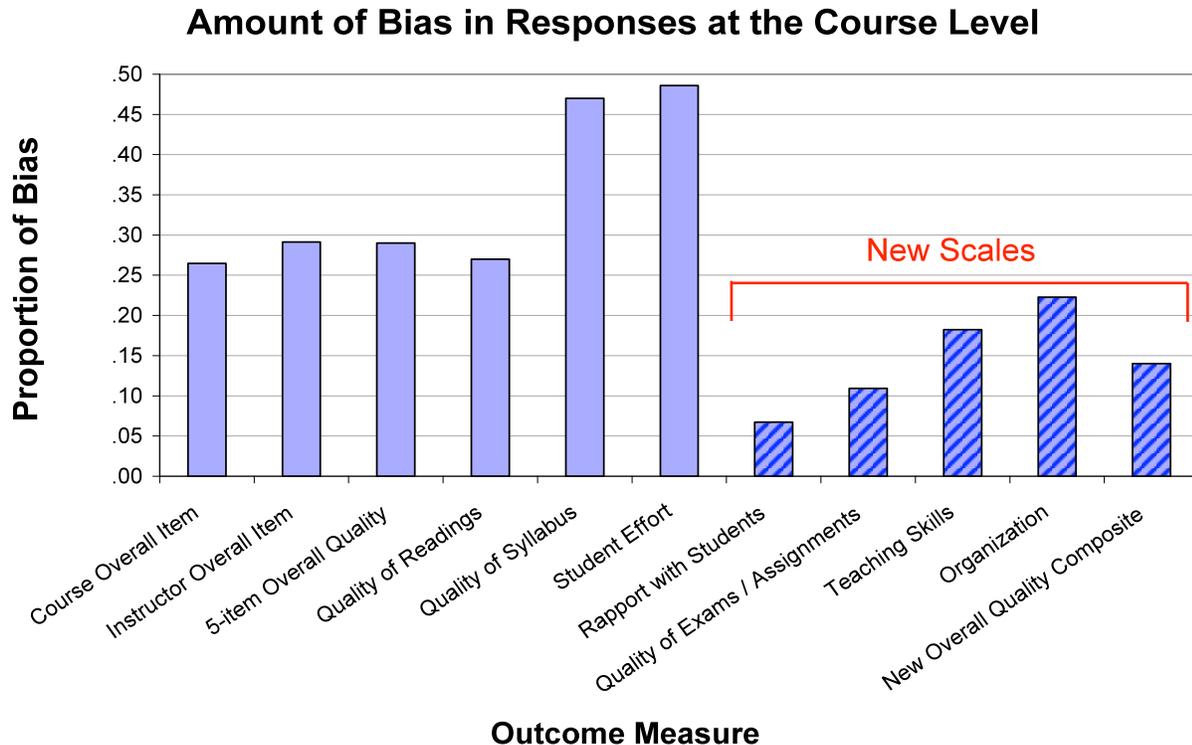
- Instructor gender
- Instructor command of spoken English
- Instructor status (tenure-track vs. non-tenure track)
- Course subject area (engineering vs. natural science vs. social science)
- Course size
- Course status (required for major vs. elective)
- Course level (100 vs. 200 or higher)
- Class time of day (morning vs. midday vs. afternoon)
- Class format (lectures vs. discussions vs. mixed)
- Student year

Although this project is truly unique in that it will ultimately be the first study in the published literature to examine such a comprehensive set of biases, there were additional sources of bias that could not be included. For example, it would have been interesting to examine possible biases associated with instructor race or more course-specific biases such as students' hostility toward race-related courses. However, given the broader goals of the project and the practical limits of what could be accomplished in a single semester, we

---

<sup>2</sup> We built 2-level HLM regression models (appropriate to the nested nature of the data) in which individual students were modeled at level 1 and courses were modeled at level 2. Sources of bias served as the predictor variables and the 2-level model allowed us to enter those predictors at both levels. For example, a predictor like instructor attractiveness could be entered at level 1 as individual students' perceptions influencing their own ratings of course quality (explaining differences in ratings between students in the same course). In addition, instructor attractiveness could be entered at level 2, allowing average ratings of attractiveness within each course to explain differences in ratings of course quality between courses. Each item on the course evaluation survey then served as the outcome variable in a separate analysis.

simply were not able to examine all possible sources of bias in this project. After identifying the final set of biases to be examined, we then determined exactly how strongly those biases influenced student responses for each of the items and dimensions identified above. The results suggested that the current evaluation questions in use at the University of Rochester were strongly influenced by these sources of bias – with estimates suggesting that 25-50% of the variability in responses was could be explained by bias (see graph). Among the sources of bias, the results suggested that three biases emerged as the strongest and most consistent predictors of student responses: higher expected grades, prior interest in the course and perceptions of the



instructor's attractiveness<sup>3</sup> were associated with higher ratings on all of the measures tested. Taken as a set, these results indicated that 25-50% of the differences in course quality ratings across the 48 courses of the study could be explained by factors like instructor attractiveness and prior interest in the course. This would mean that the average overall course or instructor ratings for a specific course could easily seem lower than another course not because there was a true difference in course quality but only because students happened to be less interested in that course prior to enrolling or because the students found that instructor to be less attractive. Estimates of 25-50% influence by these factors therefore suggest highly problematic levels of bias in the evaluation questions currently in use at the University of Rochester.

<sup>3</sup> We assessed perceptions of instructor attractiveness by having each student answer two questions: "Did you think this instructor was physically attractive?" and "Do you think other students would find this instructor physically attractive?" We chose these questions as they were simple and straightforward – hopefully minimizing differences in interpretation across students. We also put these questions near the very end of the survey so that responses to them would have the lowest chance of influencing any of the other questions on the survey.

**Goal 3: Developing a New Tool.** Given the high levels of bias associated with the current evaluation questions in use at the University of Rochester, the project sought to develop a new evaluation tool that could offer instructors more useful feedback with markedly lower levels of bias. Drawing a diverse and representative set of items from the course evaluation measures used at over a dozen universities across the globe, the project examined that item pool to: 1) estimate the degree to which responses on each item were influenced by biases, 2) identify the unique dimensions of information assessed by those items, and 3) combine that information to create a new set of questions that would offer instructors the most diverse information with the least possible amount of bias.

Gobal Questions Assessing Overall Quality. These analyses revealed that there were roughly 5 distinct sets of items in the item pool. One of the dimensions assessed overall course quality, containing the 5 items in use at the University of Rochester as well as similar items in use at other universities. However, the analyses also indicated that all of those globally worded items (e.g., What overall rating would you give this course? Rate the increase of your knowledge or skills from this course) tended to be strongly influenced by sources of bias like instructor attractiveness. This is consistent with a large body of literature on what psychologists would term global subjective evaluations. This body of research suggests that when individuals are asked to make global ratings (integrating different pieces of information into an overall score) they tend to do a poor job of integrating information and instead rely on an overall ‘gut’ feeling (or what we would call global sentiment) – causing such overall ratings to be more strongly influenced by things like the likability or attractiveness of the individual being rated.

New Dimensions. However, in addition to identifying that set of global items, the analyses revealed 4 new dimensions of information being assessed at other universities but not at the University of Rochester. The items within these 4 new dimensions tended to ask about more specific student experiences in the course (e.g., the instructor used examples in lecture that really helped me understand the material) and our analyses suggested that responses to those questions were far less influenced by sources of bias. More importantly, these 4 new dimensions could offer instructors richer and more diverse feedback on different aspects of their courses that might influence student learning and course quality. This is feedback currently unavailable to University of Rochester instructors as most of these dimensions are not currently represented in our course evaluations. For example, one dimension asks students specific questions about pedagogical skills: assessing the instructor’s abilities to summarize material, use helpful examples, fit specific topics into a larger whole and adjust the pace of lectures. Not only would average ratings on each of these 4 questions offer useful information to instructors to hone their own teaching skills, but our analyses suggest that these 4 questions are measuring a common construct and could therefore be averaged to create a “teaching skills” composite reflecting an instructor’s overall pedagogical skills. A second dimension assesses the quality of course materials: assessing the quality and clarity of exams and assignments. Not only would ratings on each of these questions provide useful information to instructors, but our analyses suggest that such questions could

be averaged to create a “quality of exams & assignments” composite to reflect the overall quality of an instructor’s course materials. The third dimension asks students questions about the instructor’s ability to connect with students: assessing the degree to which the instructor conveyed respect for students, a willingness to listen, and the degree to which the instructor made him or herself available to students. As with the previous two dimensions, not only would ratings on each of these questions provide useful information to instructors, but these questions could be averaged to create a “rapport with students” composite reflecting an instructor’s overall ability to connect with students in his or her class. The final dimension asks students questions about the instructor’s organization: assessing the degree to which the lectures were organized, the instructor was able to stay on topic during lectures, and the lectures were well prepared. Once again, not only would ratings on each of these questions provide useful information to instructors, but these questions could be averaged to create a “instructor organization” composite reflecting an instructor’s overall level of organization.

Our analyses in this project ultimately identified 13 items that were able to assess these 4 different aspects of course quality with low levels of bias:

### **1. Teaching skills**

- a. Instructor used examples in lecture / class discussions that really helped me understand the material
- b. The instructor's way of summarizing or emphasizing important points was effective
- c. As the course progressed, the instructor showed me how each of the topics fit into a whole
- d. Instructor noticed when students did not understand and adjusted the lecture pace accordingly

### **2. Quality of Exams & Assignments**

- a. The exam questions were clearly worded
- b. The exams covered important aspects of the course
- c. The assignments were helpful in understanding the material

### **3. Rapport with Students**

- a. The instructor demonstrated sincere respect for students
- b. The instructor was willing to listen to student questions and/or opinions
- c. The instructor made him or herself available for extra help

### **4. Organization**

- a. Lectures / class discussions were disorganized (reverse scored)
- b. Instructor frequently got off topic during lectures (reverse scored)
- c. The instructor tended to be ill prepared (reverse scored)

As seen in the graph quantifying bias above, responses to these sets of questions are notably less influenced by biases than the items currently being used at the University of Rochester. Thus, our results suggest that using these 13 items to assess various aspects of course quality would not only give instructors more useful feedback, but the feedback would be less influenced by factors like prior interest in the course and instructor attractiveness.

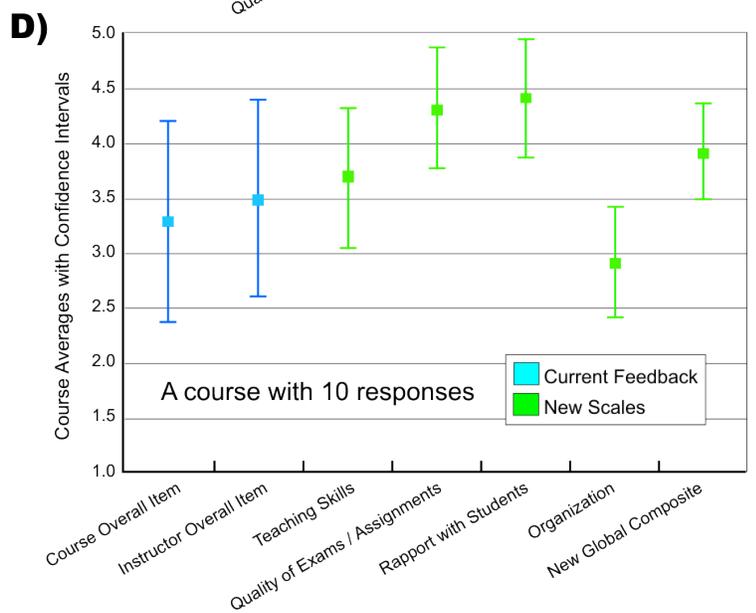
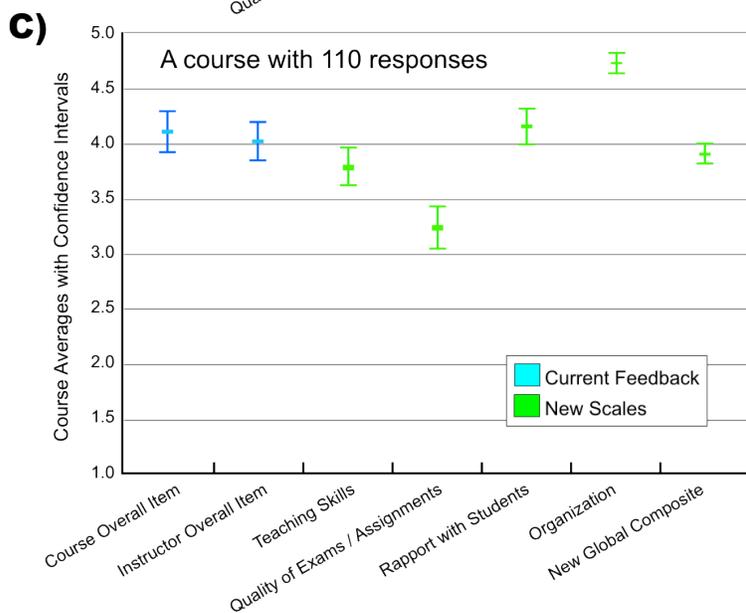
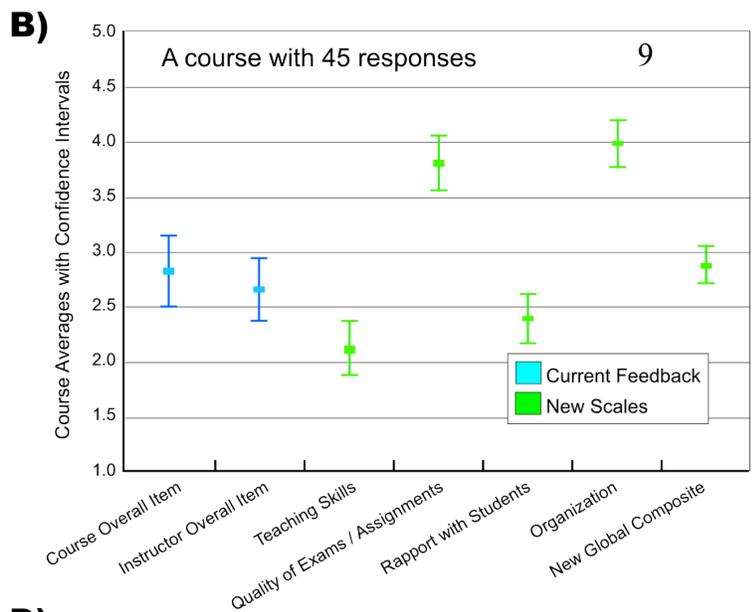
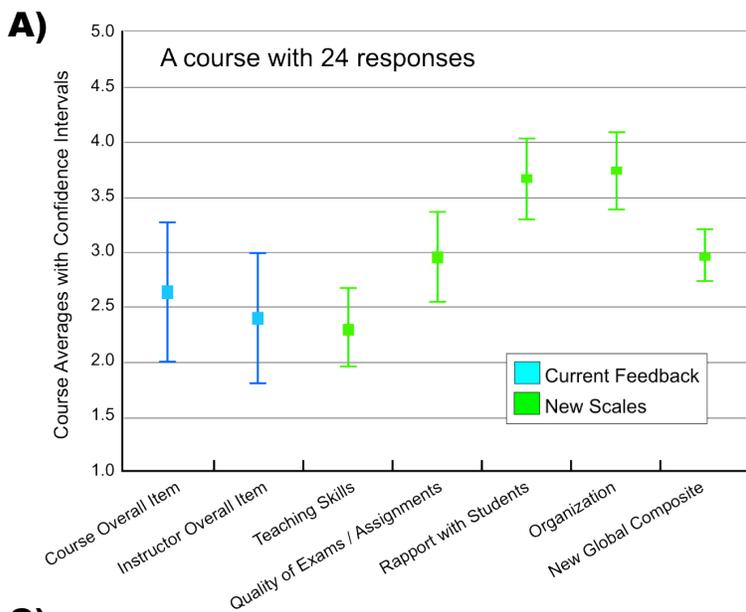
Although the items asking for global impressions tended to be excessively influenced by bias, we recognize the utility of being able to have a single global score to reflect course quality. Toward this end we

used multiple regression analyses to develop an equation to convert the four more specific dimensions described above into an overall course quality score<sup>4</sup>. Although this equation-generated overall score offers equivalent information to the two overall items currently being used, the results suggested that by using these 13 more specific (and less biased) questions to assess course quality and then combining them mathematically (rather than asking students for global ratings) the resulting information would be far less biased by factors such as instructor attractiveness, students' prior interest and expected grade. Thus, the results indicated that by switching to the 13 items above, instructors would not only get more detailed information on student's experiences in their courses (by viewing their scores on the 4 dimensions and on the individual 13 items) but could also be given an overall quality rating that is far less biased by external factors.

To help illustrate the additional information that could be obtained through the adoption of the proposed 13 items, we have created graphs illustrating the current feedback (focusing on the two items included in faculty activity reports) as well as the feedback that could have been generated from the 13 items for 4 simulated courses (based on patterns we observed in the actual course data). As seen in panel A below, the instructor teaching that simulated course would have seen an average course overall rating of 2.6 and an instructor overall rating of 2.4, suggesting lower levels of student satisfaction but failing to give specific feedback on what could be improved in the future. However, had the 13 items been part of the standard course evaluation at the University of Rochester, that instructor would have gotten a mathematically computed overall rating of 2.95 (somewhat higher once based on the 13 questions with lower levels of bias). More importantly, the instructor would have been able to see that the students believed that he or she had done a fair job of building rapport with students and remaining organized throughout the semester but found his or her teaching skills to be less satisfactory. The instructor could then have examined the average scores on the 4 items making up the teaching skills composite to get even more detailed feedback on areas that he or she could improve. Turning to panel B, the instructor teaching that course would have seen an average course overall rating of 2.8 and an instructor overall rating of 2.7 with little additional information to help guide the future improvement of that course. However, had the course evaluation been based on the proposed 13 items, the instructor would have been able to see that the students found him or her to be very organized and found the exams and assignments to be helpful but found the instructor less able to connect with the students in the course and somewhat lacking in the specific teaching skills assessed. Once again, this instructor could then have examined the average scores on the items making up those two composites to get even more detailed student feedback on areas that he or she could improve in the future. Similarly, although the 2 current items would have provided positive feedback to the instructor for panel C, the proposed 13 items could have

---

<sup>4</sup> Specifically, we developed a regression equation using scores on the 4 new composites to predict scores on a composite of global items. As a result, the mathematically derived global scores are on the same 5-point scale and correlate very strongly with global ratings, suggesting that they are providing the same information. However, as those global scores are generated mathematically from 4 scales with lower levels of bias, those global scores provide that information without the excessive bias seen in all of the global items.



informed that instructor that his or her organization was a particular strength whereas his or her exams and assignments were less satisfactory in the eyes of students. Finally, the feedback for the two current items presented for the simulated course in panel D would not only have failed to provide more detailed feedback to the instructor but those average ratings would have had high levels of error when based on so few responses (based on the confidence intervals suggesting that the means were estimated with an error of  $\pm 0.8$  points in this simulated course). Had the feedback been based on the proposed 13 items, the instructor would have gotten a more accurate estimate of overall quality (with errors of only  $\pm 0.4$  points) and would have seen that the students found him or her to be somewhat disorganized. Thus, by adopting the proposed 13 items and giving instructors feedback on 1) the new global composite, 2) the average ratings on the 4 new dimensions, and 3) the average ratings on the individual items, the University of Rochester would be providing instructors valuable information to hone the quality of their courses and would be providing that information with greater

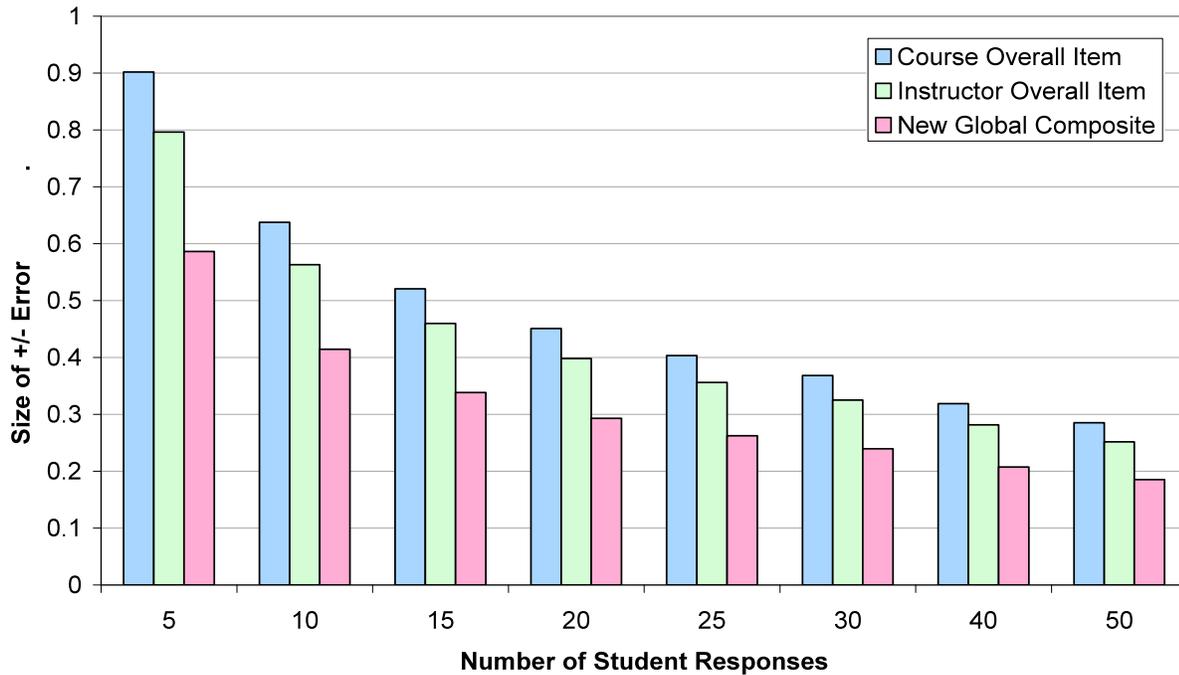
levels of accuracy (lower error) and lower levels of bias by factors like previous interest in course, expected grade and instructor attractiveness.

**Goal 4: Contrasting Online to In-Class Administration.** Using the same modeling approach used in Goals 2 and 3, we built models allowing the sources of bias assessed in the project to predict responses on each of the current evaluation questions as well as on the 13 proposed questions of the new scale. In these models, we introduced terms to determine if: 1) average ratings on each item were higher or lower with online-administration, and 2) if any of the sources of bias had stronger or weaker effects with online-administration. This multi-level multivariate approach offered a powerful method of detecting possible degradation of information from online administration. However, the analyses failed to identify any average differences between evaluations collected online vs. in-class on any of the items or composites examined, suggesting that average course ratings might in fact be comparable across the two methods of administration. Furthermore, when method of assessment (online vs. in-class) was introduced into the models, it only accounted for a small amount (0-4%) of the variability in course ratings. This suggests that 96-100% of the differences in course ratings observed between courses was completely unrelated to *how* those ratings were obtained. Given the diverse array of courses involved in the project, the rigorous design of the project and the large number of student responses supporting these analyses, the lack of significant findings for online administration biases is actually quite striking. Taken as a set, these results suggested that collecting course evaluation data online does not seem to adversely impact the quality of information obtained despite the markedly lower response rates associated with that method.

**Goal 5: Determining Number of Responses Needed.** Finally, given the lower response rates associated with online administration of course evaluations, the project directly examined the levels of uncertainty that arise with exceedingly small numbers of respondents (i.e., as few as 5 student responses per class). Using the 88 responses from one of the larger courses in the study as a basis, we calculated the 95% confidence intervals for class averages based on subsamples of 5, 10, 15, 20, 25, 30, 40 and 50 students. Confidence intervals provide upper and lower boundaries for a sample average (e.g., average rating on a course evaluation question), typically presented as plus or minus a certain amount of points on the scale to indicate the error associated with that estimate. As seen in the figure below, the two overall quality items demonstrated relatively high levels of error when estimated in very small (e.g., 5-10) samples of students. In fact, had the average for the item, “What overall rating would you give this course?” been based on only 5 student responses, that mean would only have been accurate to only +/- 0.9 points. Thus, had the instructor for this course gotten an average of 3.5 on that item from only 5 responses, the true mean for his course could have been as high as 4.4 or as low as 2.6. That reflects a high level of uncertainty, and suggests that class averages based on exceedingly small subsamples of responses (e.g., as few as 5 or 10 responses from courses with far greater numbers of students) should be interpreted very cautiously. The graph also

indicates that as response rates increase to samples of 40-50 students, these error rates drop rapidly. Thus, for larger courses (e.g., courses with 80 or more students), the results suggest that even responses from as few as 40-50 students might still provide reasonably accurate estimates of course quality<sup>5</sup>. Turning to the results for

### 95% Confidence Intervals by Number of Responses



specific course evaluations items, the results indicated that it would require a minimum of 40-50 responses to obtain averages on the two overall items currently in use at the University of Rochester that were accurate to +/- 0.3 points on the 5-point scale. In contrast, the global composite based on responses to the 13 proposed questions would only require 20 responses to obtain a similar level of accuracy. Although this is not surprising as multi-item scales tend to give more reliable and accurate information, the graph demonstrates just how much accuracy is gained by switching to a mathematical composite in lieu of the two global items – literally offering comparable levels of accuracy with half as many responses (20 vs. 40 responses to obtain an accuracy of +/- 0.3 points). Thus, the results of the project not only suggest that the new 13-item scale would provide more diverse and less biased information, but they also suggest that the global composite based on those 13 items would offer higher levels of accuracy – particularly in courses with low participation rates. It is worth noting that for large courses (e.g., over 100 students) these estimates of error are relatively

<sup>5</sup> This argument primarily applies to larger courses. If a course has only 20 students and 18 of them provide responses on the course evaluation, despite the uncertainty arising from means based on only 18 responses the resulting means could still be considered a fair representation of the course quality as almost all of the students would have participated. Similarly, 7 responses in a course of 8 students would also give a fair representation of the course quality. However, in moderate to large courses (e.g., with 40 or more students enrolled), means based on only 18 or 7 responses would be far more circumspect and the uncertainty would be of greater concern.

independent of course size. Thus, even in a course of 200 students, obtaining 50 responses would give relatively low error rates – particularly on the proposed global composite based on the 13 items. This provides additional evidence to suggest that the lower response rates observed from online administration might not be problematic, particularly for larger courses where even a 30% response rate would yield 50 or more student responses.

### **Project Recommendations**

- 1) **Discontinuing use of excessively biased items.** The results strongly suggested that the two questions currently representing quality of instruction (overall ratings of course and instructor) as well as a larger composite of the 5 course quality questions currently in use were strongly influenced by sources of bias. Based on these results, we would suggest that the university consider discontinuing use of such global items. As the students' ratings of their own effort in the course also demonstrated excessively high levels of influence by bias, we would suggest those be dropped from course evaluations as well.
  
- 2) **Shifting to more specific domains.** The results identified 4 distinct dimensions of course quality (teaching skills, quality of exams/assignments, rapport with students, organization) assessed at other universities. We would recommend that the University of Rochester consider shifting the focus of our course evaluations to these dimensions (instead of simply using global evaluative questions) as it would provide instructors more detailed and useful feedback on their courses with lower levels of bias.
  
- 3) **Using multiple items to assess each domain.** The results suggested that the use of individual items to assess quality (as is the current practice) leads to problematic levels of error when small numbers of student responses are obtained for a specific course. Specifically, when a course average is based on a subset of only 10 students providing course evaluations, the averages on either of the overall items have error rates of roughly +/- 0.6 points, suggesting high levels of inaccuracy. As online administration has led to markedly lower response rates, this is concerning because classes with only 25 students could easily have 10 or fewer students provide course evaluation data. However, the data also indicated that by using 3-4 items to assess a domain (rather than 1) and by creating a global quality composite based on all 13 items, it was possible obtain estimates with higher accuracy (lower error), even in smaller samples. Consequently, we identified 13 items (with notably lower levels of influence by bias) to assess the four proposed dimensions. We would recommend the University of Rochester consider adopting the use of these items for assessing course quality in lieu of the global items currently in use.
  
- 4) **Using an equation to generate global-quality scores.** As the results suggested that asking students globally evaluative questions on course quality introduced high levels of bias, we would recommend

using an equation to synthesize the 4 proposed dimensions into an overall score. This equation was developed in the current dataset to most closely represent the information obtained by global questions (e.g., What is your overall rating of this course?) but offers that information with markedly lower levels of bias and higher levels of accuracy. In fact, the results suggest that it provides an estimate of overall quality that is relatively stable, even when based on as few as 15-20 responses.

- 5) **Discontinuing use of largely unused questions.** The current course evaluation survey offers an open-ended question asking for comments after each and every numeric question. This creates an additional 16 items on the survey. However, less than 2% of students use the majority of those open-ended questions. Most students constrain their open-ended comments to the strengths and weaknesses questions regarding the instructor and the strengths and weaknesses questions regarding the course. In the interest of parsimony, we would recommend simply retaining those open-ended questions that the students actually use.
- 6) **Limiting interpretation of means from small samples.** Given the higher levels of inaccuracy resulting from means calculated in samples of fewer than 20 respondents, we would recommend high levels of caution when interpreting course data from smaller numbers of responses – particularly when that represents only a fraction of the students in the course.
- 7) **Presenting confidence intervals with averages.** To more directly address the issue of error, all instructors could be given 95% confidence intervals for each of the averages presented to them in their student evaluation feedback. The equations for this are very straightforward and could easily be programmed into the standard online feedback. Providing this information would enable the instructors to determine for themselves the level of healthy skepticism appropriate for any set of means – particularly for means generated from smaller numbers of student responses.

## CURRENT COURSE EVALUATION QUESTIONS

Question	Select a response below						
Status of this course in your program?	Major <input type="checkbox"/>		Elective <input type="radio"/>		Other Requirement <input type="radio"/>		Uncertain <input type="radio"/>
Class Year?	Freshman <input type="radio"/>	Sophomore <input type="radio"/>	Junior <input type="radio"/>	Senior <input type="radio"/>	Graduate <input type="radio"/>	Non Matriculated <input type="radio"/>	Other <input type="radio"/>

### On the student

Question	Select a response below				
Rate the level of your involvement in the activities of this course (for example: attendance, participation, completing assignments).	Fully Engaged <input type="radio"/>	Mostly Engaged <input type="radio"/>	Average <input type="radio"/>	Partially Engaged <input type="radio"/>	Minimally Engaged <input type="radio"/>
Comments:					
Rate the increase of your knowledge or skills from this course.	Greatly Increased <input type="radio"/>	Increased <input type="radio"/>	Average <input type="radio"/>	Minimally Increased <input type="radio"/>	Not Increased <input type="radio"/>
Comments:					
What overall rating would you give yourself as a student in this course?	Very Hardworking & Professional <input type="radio"/>		OK <input type="radio"/>	Very Lax & Unprofessional <input type="radio"/>	
Comments:					

### On the course

Question	Select a response below				
How well did the syllabus describe the course content?	Extremely Well <input type="radio"/>	Very Well <input type="radio"/>	Well <input type="radio"/>	Not Well <input type="radio"/>	N/A <input type="radio"/>
Comments:					

The readings were important in my learning in this course.	Very Important <input type="radio"/>	Somewhat Important <input type="radio"/>	Not Important <input type="radio"/>	N/A <input type="radio"/>	
Comments:					
How well did the course assignments and exams support course objectives?	Extremely Well <input type="radio"/>	Very Well <input type="radio"/>	Well <input type="radio"/>	Not Very Well <input type="radio"/>	Poorly <input type="radio"/>
Comments:					
What are the major strengths of this course?	STRENGTHS:				
What are the major weaknesses of this course? Please make suggestions for improvement.	POSSIBLE IMPROVEMENTS:				
What overall rating would you give this course?	Excellent <input type="radio"/>	Very Good <input type="radio"/>	Average <input type="radio"/>	Not Very Good <input type="radio"/>	Very Poor <input type="radio"/>
OVERALL Comments on COURSE:					

### On the Instructor

Question	Select a response below				
How responsive was the instructor in and out of class?	Very Responsive <input type="radio"/>	Mostly Responsive <input type="radio"/>	Average <input type="radio"/>	Minimally Responsive <input type="radio"/>	Unresponsive <input type="radio"/>
Comments:					

How effective was the instructor's teaching in this course?	Very Effective <input type="radio"/>	Effective <input type="radio"/>	Average <input type="radio"/>	Minimally Effective <input type="radio"/>	Ineffective <input type="radio"/>
Comments:					
I have a stronger interest in this subject because of the instructor.	Strongly Agree <input type="radio"/>	Agree <input type="radio"/>	Neutral / Mixed <input type="radio"/>	Disagree <input type="radio"/>	Strongly Disagree <input type="radio"/>
Comments:					
What are the major strengths of this instructor?	STRENGTHS:				
What are the major weaknesses of this instructor? Please make suggestions for improvement.	POSSIBLE IMPROVEMENTS:				
What overall rating would you give this instructor?	Excellent <input type="radio"/>	Very Good <input type="radio"/>	Average <input type="radio"/>	Not Very Good <input type="radio"/>	Very Poor <input type="radio"/>
OVERALL comments on INSTRUCTOR:					

If there are any other further comments you would like to make about this course, please do so in the space provided below.

Additional Comments on INSTRUCTOR:

## PROPOSED COURSE EVALUATION QUESTIONS

In this class...	Not at all	A little	Somewhat	Quite a bit	Very
Instructor used examples in lecture / class discussions that really helped me understand the material	<input type="radio"/>				
The instructor's way of summarizing or emphasizing important points was effective	<input type="radio"/>				
As the course progressed, the instructor showed me how each of the topics fit into a whole	<input type="radio"/>				
Instructor noticed when students did not understand and adjusted the lecture pace accordingly	<input type="radio"/>				
The exam questions were clearly worded	<input type="radio"/>				
The exams covered important aspects of the course	<input type="radio"/>				
The assignments were helpful in understanding the material	<input type="radio"/>				
The instructor demonstrated sincere respect for students	<input type="radio"/>				
The instructor was willing to listen to student questions and/or opinions	<input type="radio"/>				
The instructor made him or herself available for extra help	<input type="radio"/>				
Lectures / class discussions were disorganized	<input type="radio"/>				
Instructor frequently got off topic during lectures	<input type="radio"/>				
The instructor tended to be ill prepared	<input type="radio"/>				

What are the major strengths of this instructor?	INSTRUCTOR STRENGTHS:
Please share any suggestions for improvement.	POSSIBLE AREAS FOR IMPROVEMENT:

What are the major strengths of this course?	COURSE STRENGTHS:
Please share any suggestions for improvement.	POSSIBLE AREAS FOR IMPROVEMENT:

ADDITIONAL COMMENTS:
----------------------