

A university-wide system of student ratings of teaching effectiveness

Academic Affairs Committee Student Ratings Report

ROCHESTER INSTITUTE OF TECHNOLOGY

March 21, 2013

Authored by: AAC at RIT

Report

Academic Affairs Task Force on a University-Wide System of Student Ratings of Teaching Effectiveness

Academic Affairs Committee (AAC)

Roxanne Canosa (GCCIS)*

Bill Evans (SCB)

Carol DeFilippo (NTID)*

Liz Kmiecinski (CHST)

Chris Licata (Delegate of Jeremy Haefner, ex officio, voting)*

David Long (CIAS)

Jeff Lodge (COS)

Frank Sciremammano (KGCOE)

Linda Tolan (CAST)

Tracy Worrell (CLA)*

Mike Laver (CLA, At Large)*

Ed Mineck (NTID, At Large)

Marla Schweppe (CIAS, At Large)

Supplemental Task Force Members

David Hostetter (ITS)*

Karel Shapiro (Senior Staff Specialist)*

*AAC subcommittee members responsible for preparing this report

March 21, 2013

Table of Contents	Page
I. Executive Summary	3-4
II. Prologue	5-6
III. Background	6-10
A. 2009-2010	
B. 2010-2011	
C. 2011-2012	
1. Four key conclusions	
2. Recommended action step	
D. 2012-2013 Pilot Study	
1. Pilot Design	
2. Participants	
3. Results	
a. Feedback from students	
b. Feedback from faculty	
c. Feedback from unit heads	
d. Feedback from deans	
IV. Recommendation from Academic Affairs Committee 2013	11-14
A. Recommendation #1	
B. Recommendation #2	
C. Recommendation #3	
D. Recommendation #4	
V. Survey Administration	14
A. Personnel	
B. Procedures	
VI. Establishment of an Effective System of Student Ratings	14
References	15-22
Appendix A: Rationale for seven core items for student ratings of teaching effectiveness	
Appendix B: Student and faculty pilot survey for feedback on SmartEvals and IDEA Center	
Appendix C: Responsibilities as listed by SmartEvals	
Appendix D: Student ratings price quote - SmartEvals	

I. EXECUTIVE SUMMARY

The Academic Affairs Committee (AAC) was charged by the Executive Committee of Academic Senate to “**conduct a pilot investigation of two systems of student input on teaching effectiveness in order to determine a final recommendation of a system for university-wide launch in fall semester, 2013.**” This charge is an extension of the work done last academic year in which the AAC conducted an extensive review of the literature and best practices surrounding online student ratings of teaching effectiveness, evaluated a number of potential vendors, and finally determined that two vendors in particular could suit the needs of the RIT community at this particular time. That work was embodied in an extensive report to Academic Senate, at which time senate voted to charge this committee to conduct pilots of the two vendors, and evaluate the results so as to recommend a campus wide system to be implemented at RIT in fall semester of academic year 2013-2014.

This committee would be remiss if it did not acknowledge the tremendous amount of work done by previous committees and taskforces, and so to that end we would like to acknowledge in particular the Academic Affairs Committee of 2010-2011 as well as the Course Evaluation Task Force of 2009-2010. This committee’s work is very much built on the solid foundations that these previous bodies have established.

In fulfillment of this year’s charge, the Academic Affairs Committee has conducted an extensive pilot study of two vendors: SmartEvals and IDEA Center. These two vendors were chosen because they represented two very different approaches to student ratings of teaching effectiveness: with SmartEvals, RIT is responsible for providing its own core items while IDEA provides a predetermined set of items and in return gives faculty an individualized “diagnostic” report as to their teaching effectiveness. In both cases individual faculty members, departments and colleges can add items, although that number is limited to 20 in the IDEA system, whereas in SmartEvals, that number is theoretically unlimited. The core items for the SmartEvals system that the AAC arrived at after extensive review of RIT faculty feedback as well as a review of the literature to determine best practices are:

1. The instructor enhanced my interest in this subject.
2. The instructor presented the course material in an organized manner.
3. The instructor communicated the course material clearly.
4. The instructor established a positive learning environment.
5. The instructor provided helpful feedback about my work in this course.
6. The instructor supported my progress towards achieving the course objectives.
7. Overall this instructor was an effective teacher.

(Scale: 1=Strongly Disagree to 5=Strongly Agree)

8. I attended this class regularly. No Yes

In addition, two open-ended questions are also recommended:

- “What did this instructor do well?” and
- “How can this instructor improve?”

The pilot of both systems was conducted in the fall quarter of this academic year. The quantitative and qualitative data garnered from student and faculty participants, as well as feedback from academic unit heads and deans was collected and analyzed over the course of the winter quarter so as to be able to present the findings to academic senate in the spring quarter. The conclusion that the committee reached is that both students and faculty seemed to respond more favorably to the SmartEvals system than to the IDEA system. While there are many reasons for that outcome (as found on pages 8-10 in the report), a major reason that came up in every constituency the AAC consulted was that the IDEA system, while impressive in its diagnostic capabilities, is simply too long at almost fifty core items. There is a real fear that students would quickly succumb to survey fatigue. On the other hand, the array of benefits that SmartEvals would bring to the RIT community is considerable, while at the same time allowing for a core set of items that is much shorter than the IDEA system.

In conclusion, the formal motion that the AAC presents to Academic Senate is:

The Academic Senate endorses the report of the Academic Affairs Committee concerning online student ratings of teaching effectiveness, including the following four recommendations:

- 1. Use the SmartEvals system to gather student ratings of teaching effectiveness in classroom settings across the university.**
- 2. Use the same set of established core items across the university that were used in the pilot ($\alpha = .93$ from pilot).**
- 3. Provide the online results for an individual instructor (except for instructor added items) only to the instructor, instructor's immediate supervisor and dean, the provost, and tenure and promotion committees per college guidelines.**
- 4. Re-evaluate recommendations 1-3 after three years of data collection with SmartEvals.**

In addition to these four recommendations, there is also a second, supplemental report that will address aspects of implementing an online student ratings system according to best practices as determined through this committee's extensive research. Such topics include pre-launch communications to the RIT community, strategies to encourage survey participation, data analysis and reporting, and a host of other relevant topics.

II. PROLOGUE

As we look forward to the launch of our first university-wide system of student ratings, we reflect on where we came from and where we are now in order to recognize the import of the move we are about to make. Given the comprehensive and far reaching changes that RIT finds itself in the midst of, the natural question is, Why change? Why now?

Up to now...

Appendix A in the report of the 2010 Task Force documented that there are as many student ratings schemes on campus as there are colleges, and almost as many practices (or lack of practice) in the uses of student ratings as there are academic units. Whether we approach ratings with the motivation of student or instructor or faculty supervisor, such inconsistencies can hamper achieving the desired outcome or engendering any sense of equity across the university. This outcome is to provide a student rating system that contains a set of core items consistently used by every college which also provides sufficient flexibility for colleges, programs and faculty to add other items of interest.

Further, it has been one of the initiatives of our Provost to lead us toward a more balanced view of student ratings. In several communications, the Provost has promoted a shift toward using multiple types of evidence in evaluating each faculty member's responsibility. No longer can student ratings alone determine thumbs up or down in the area of teaching effectiveness, and nor should they, given best practices and several decades of research in this area.

In addition, in those parts of the university where student ratings rely on OCE (Online Course Evaluations) or SCANTRON (optical scanning system for paper forms), the limits of technology compel change. Neither system is physically capable of managing the volume of surveys we will generate as a university, nor can they match the functionality that exists in many of the online student ratings vendors that are increasingly used among our peer institutions.

Now...

To our knowledge, while all of the ratings surveys in use at RIT provide data about an individual's teaching effectiveness, none has been subjected to external validation. Statistics have been generated on the reliability and internal validity of some (though not all) units' survey items, but it has not been shown that any constitute a measure of teaching effectiveness. Similarly, the charge to this committee was limited to the development of a reliable and internally valid tool. To have validated this tool regarding the construct of teaching effectiveness would have required a major investment of time, well beyond our two-year limit, as well as monetary resources and the trust of our colleagues, as such a study presumes access to confidential data that typically is not granted to a university faculty committee.

Instead, we relied on research-based published reports about teaching effectiveness, the dimensions of teaching that students can reliably provide input on, how to construct a useful student ratings survey as well as extensive analysis of literature surrounding student ratings. Our selection of the core items was informed by this literature (summarized in our report of May 2012). The recommended core items address dimensions of teaching effectiveness that students can observe and behaviors that they can rate, based on their classroom experience with an

instructor. The items and the procedures we recommend follow best practices. Our pilot data confirm a satisfactory level of reliability and internal validity such that we recommend our list of core items with confidence.

In the future...

Have we learned all that we need to know about the recommended ratings survey in one pilot administration? The answer is no. Two areas, in particular, will require careful monitoring:

First, we are aware that response rate is a concern of many faculty, a concern that is intensified in online ratings systems. In fact, our choice of vendor was strongly influenced by a desire to have a short survey in order to avoid jeopardizing response rates. Nonetheless, we expect that initial online response rates will be disappointing to some faculty, especially compared to rates that can be obtained with an in-class paper survey. For this reason, we have recommended that the campus coordinator and college liaisons of the system promote and facilitate use among the faculty of known strategies that are easy to implement and that have been shown to yield moderately high response rates.

Second, we are aware that there are individual faculty members who are at or near a critical transition point in their academic career, whether that be pre-tenure review, tenure, or promotion. Student ratings will comprise an important segment of the evidence to be examined and we want to ensure that a new student rating system does not have a negative affect on any such personnel decisions. To this end, in addition to the best practice procedures we are recommending, we have outlined a research plan that can identify any problems associated with a change in survey. In addition, each college has the capability to add unique items from their current survey to the new university core items in order to collect specific comparison data. In the end, however, the literature fairly uniformly suggests that there is no non-response bias in online student ratings systems, and if any quantitative change occurs, an individual faculty member's ratings are likely to be slightly higher in an online scheme.

Therefore...

We acknowledge and respect the desire to keep to familiar paths, and the reluctance to embark in a new direction, leaving the known behind. In deference to (a) the need for consistency and equity in adherence to policy on faculty evaluation across the university, and given (b) strong indications that we have crafted a useful set of reliable and valid core items in agreement with well-researched descriptions of instructional effectiveness, and with (c) strategies in place to monitor issues of possible concern, we ask our colleagues to consider the recommendations that follow.

III. BACKGROUND

A. 2009-2010

The Provost in partnership with Academic Senate charged a task force (Course Evaluation Task Force) to study the current methodologies used at RIT to solicit formal student feedback on instruction and offer specific recommendations for:

- A consistent university-wide **on-line** process for student ratings of teaching;

- A set of 5-7 **core items** that adequately reflect the essentials of teaching effectiveness and which meet a sufficiently high level of validity and reliability so as to serve as one component of the annual review and be constructed and administered within the procedural guidelines established by research and best practice in the field;
- A bank of **customizable items** that would serve a formative use

The Task Force submitted its report and recommendations to the Provost and Academic Senate in August 2010.

B. 2010-2011

Following discussion related to the Task Force Recommendations on the floor of Academic Senate, the Academic Senate determined that further community input and deliberations were needed. As a result, the Senate charged the Academic Affairs Committee to oversee a thorough vetting process in the colleges of the 2010 report and recommendations. Through the administration of a comprehensive faculty survey, considerable issues and concerns were raised by faculty with respect to certain of the recommendations. In particular, faculty questioned the relevance, ramifications and focus of proposed core survey items and web posting of aggregate results for viewing by RIT community members.

These concerns prompted the Academic Senate to charge the Academic Affairs Committee with further analysis during the 2011-2012 academic year and with the preparation of a new set of recommendations which would address faculty issues and concerns raised and respond to the original charge. In carrying out this work, the committee was asked to revisit best practices in the field in order to:

- Establish a set of principles to inform the development of an online student rating system and
- Identify guidelines by which such a system would be implemented at RIT

C. 2011-2012

The Academic Affairs Committee concluded its work and submitted its report on a “University-Wide System of Student Input on Teaching Effectiveness” to Academic Senate (May 2012). The report outlined four key conclusions and a recommended action item.

Four Key Conclusions:

1. An extensive literature review which built on the findings outlined by the Task Force on Course Evaluations (2010) confirmed that:
 - a. Student rating of instruction is both reliable and valid as an indicator of teaching effectiveness;
 - b. Students prefer online methodology and will provide much more qualitative data online than they do in traditional paper format.
 - c. Further, research reveals that instructor ratings from students do not vary in a statistically significant way in online versus traditional paper surveys.
 - d. Non-response bias (which tends to be a factor in faculty resistance to the online format,) is not borne out in the research—although research does suggest that

response rates may be lower in online surveys, particularly in the first few years and that measures may have to be adopted to try to boost the rates of response.

2. A set of principles and guidelines based on research and best practice, steered the analysis of the design and functionality of student rating systems available in the market. These overarching principles produced 11 potential, desired “product” characteristics:

- a. The on-line student rating tool must provide both summative and formative feedback
- b. The rating tool should focus on providing feedback only on areas students can judge
- c. The tool should adjust for variables affected by student characteristics
- d. The system will guarantee student anonymity
- e. The system will preserve the confidential nature of the faculty appraisal process
- f. The rating tool will have formative features that provide diagnostic information for improvement
- g. The rating tool will be flexible to accommodate varieties of courses and teaching methods
- h. The rating tool will incorporate the necessary analyses to account for known effects
- i. Students will be educated about the purposes and uses of survey results
- j. Results will be interpreted as one source of information and evidence for teaching effectiveness
- k. Results will be used in the annual review and to foster ongoing professional development and improvement

3. Following in-depth interviews with representatives of those vendors that appeared to meet the functional and technical requirements outlined for RIT’s on-line rating system, two vendors, in particular, emerged as the strongest contenders: SmartEvals and the IDEA Center.

4. A set of core items needed to be developed for the SmartEvals survey. The IDEA survey provided their own items which could not be altered. While the core items recommended by the first task force were used as a starting point, these items were revised given faculty feedback and further research from the field (See Appendix A).

Recommended Action Step:

The Academic Affairs Committee recommended to Academic Senate that RIT “conduct a pilot investigation of two systems of student input on teaching effectiveness in order to determine a final recommendation of a system for university-wide launch in fall semester, 2013.” The Academic Senate endorsed this recommendation on May 10, 2012.

While this action step delayed the implementation of a new student rating system by one year, the Academic Affairs Committee believed that it was necessary to observe how our instructors and students interacted with each system in order to determine which one better suited our campus needs. *The committee recommended that the campus examine faculty and student*

experiences with both IDEA Center and SmartEvals before making a large-scale commitment to either system.

This pilot would give the community a taste of the capabilities of both products and would allow the Committee to gather feedback from faculty, students, unit heads and deans about both products. The committee recognized as did Academic Senate that the two tools were different enough in terms of survey length, cost, faculty effort required to set up the survey, digital reporting capabilities, and system support for both summative and formative purposes that gaining some first-hand experience with both would help inform the decision. Both vendors indicated that they would provide the survey tools at no cost to RIT during the pilot period.

D. 2012-2013: Pilot Study

A pilot was conducted during fall 2012 to gauge each system's suitability to RIT by gathering student, faculty, unit heads and deans' feedback.

1. Pilot Design

A representative sample of faculty from each college participated in the pilot. Faculty were eligible if they were tenured, tenure-track in final pre-tenure year, or lecturers. In order to give faculty consistency in using both systems, the pilot was limited to those faculty teaching two sections of the same course or one large section of a course that could be split between the two vendors. Students in a class used one of the rating tools and in addition were asked to complete a follow up survey to determine their opinions regarding the specific tool. Faculty were also asked to complete an opinion survey after receiving and reviewing their reports from each system vendor (see Appendix B).

2. Participants

- 58 faculty from across the university used both surveys in sections of the same course
- 128 class sections were used
 - Total number of students as possible participants; SmartEvals=1421 students; IDEA = 1524 students
- Response rate for Student Ratings Survey:
 - SmartEvals: 59%
 - IDEA: 51%
- Response rate for post-ratings/feedback survey on each vendor.
 - Faculty: 68%
 - Students: SmartEvals = 28% (238); IDEA = 43% (337)

Focus groups were also conducted to gather feedback from Unit heads (20 representatives from all colleges on campus) and Deans (representatives from every college).

3. Results

a. Feedback from Students:

- Quantitatively: Students reported more overall positive feelings (scale $\alpha = .80$) towards the SmartEvals survey ($M = 4.05$) than the IDEA survey ($M = 3.91$), $t(544) = 2.537, p = .011$. Individual items showed little significant difference except for the survey being too long and the survey being easy to fill out. (SmartEvals, $M = 2.27, M = 4.26$; IDEA, $M = 2.87, M = 4.11$; respectively)
 - Qualitatively: *What topics/questions related to teaching effectiveness were overlooked with this survey, if any?* **SmartEvals:** None (27), workload, teacher personality, teaching style, too few items, availability, teacher communication; **IDEA:** None (76), items not applicable to all courses, style of teaching, more qualitative space, content value, personality, availability, fairness. *What additional comments do you have about this online system of student feedback on teaching effectiveness?* **SmartEvals:** None (16), easy to use, quick, thorough, needs more course items, more items needed, prefer old system, anonymity, usability; **IDEA:** None (44), surveys not needed, quick, easy, efficient, some items unnecessary, confusing, usability, too generic, too long, old system is fine.
- b. **Feedback from Faculty:**
- Quantitative: Faculty reported more overall positive feelings (scale $\alpha = .88$) towards the SmartEvals survey ($M = 3.52$) than the IDEA survey ($M = 3.19$), however the difference was not significant. Individual items showed little significant difference except for the item “I understood how to interpret my ratings report. (SmartEvals, $M = 3.94$; IDEA, $M = 3.26$).
 - Qualitatively: *Please let us know why you have chosen this product.* **SmartEvals:** simple, fewer items, easy to add items and see online feedback, user-friendly, clear easy output, good suggestions to increase response rates, appears to provide better information, better response rate, less confusing (than IDEA), better than current system, fast feedback, good communication with users, liked online report, IDEA has too many items and too complex of a report; **IDEA:** intuitive, easy to grasp, robust instrument, effective for reflecting upon goals for the course, easy to interpret results and easier to set up, teaching rated according to goals, more comprehensive, items for “important & essential objectives” helpful, relevant details given, straightforward, new information and way to look at the data, SmartEvals is less helpful in terms of improving the course in future years.
- c. **Feedback from Unit heads:**
- Qualitative: Both systems could fit the needs of the unit heads present, however, IDEA was deemed too long and of concern to most unit heads. Unit heads preferred the customizability of SmartEvals, however, they liked the comparability, professional development focus, and reliability/validity of IDEA.
- d. **Feedback from Deans:**
- Qualitative: Deans were also concerned with the length of IDEA’s survey, possible student fatigue responding to five IDEA surveys each term. Deans reported only needing a small set of items to assess an instructor’s overall teaching effectiveness. Deans would like a reliable system with potential to compare to other institutions. Response rates were of large concern. Questions were also raised about hosting our own data and security of each system.

IV. RECOMMENDATION FROM ACADEMIC AFFAIRS COMMITTEE 2013

Given the original charge (2009-2010), pilot study feedback results from students, faculty, unit heads and deans, the current teaching/learning culture at RIT, and the overarching goal to adopt a system that encourages high response rates, the committee makes the following four recommendations:

Recommendation #1: Use the SmartEvals system to gather student ratings of teaching effectiveness in classroom settings across the university.

Rationale: Why SmartEvals over IDEA Center

	SmartEvals	IDEA Center
Familiarity	Information more like faculty and admin. are used to	Report looks complicated and takes time to interpret
Simplicity	Limited set of core items with no action from faculty needed	Benefits from diagnostic report depend on faculty form
Speed	Short survey for students	Long item-set (47) expected to burden students, rater fatigue
Flexibility	Brief, so items added need not be onerous	No flexibility to core items set, limit to added items
Completion	Fewer items favors completion of entire survey	Concern of dropout rate due to length of survey
Response Rates	Brevity and email tips should favor higher responses	Concern of rate decrease across years due to length
Program Needs	Core items don't address objectives, avoids possible conflicts	Concern of specifying objectives at instructor level and possible confusion
Reporting	Timely, web-based reporting allowing for customization	Longer distribution of reports via .pdf
Cost	Low cost	Higher cost for fully loaded system that may not be utilized

Additional advantages of SmartEvals:

- Provides a “drill-down” capability allowing instructors to view results for selected subsets of the data.
- Offers suggestions of formative items from a bank of items used by its customers. These are available for suggestions and possible comparisons to other institutions.
- Allows creation of a faculty action plan that will provide guidance to faculty about how to enhance instructional effectiveness.

Other Specific benefits of SmartEvals to the RIT Community

- **Students**
 - Access a personal web page that presents rating forms for courses.
 - Assurance of anonymity of their responses (easy integration with Shibboleth so students login with RIT username and password).
 - Paperless system.
 - Notification via e-mail when rating periods open and close.
 - Notifications to non-responders about courses to be completed.
 - Ability to complete ratings on a Smartphone.
- **Instructors**
 - Can add additional items to the core set.
 - Receive summary statistics for each of the core items and for all items that are added (mean, standard deviation; response rate).
 - See their average scores compared to averages at the department, college and university levels.
 - Receive unedited student responses to open ended comments.
 - Can access their historical ratings data from past terms.
 - Export reports in a variety of formats (i.e. Excel, .pdf).
- **Unit heads**
 - Able to customize a set of items to be added for all faculty at the department or course level.
 - View same core information as the faculty member.
 - Able to set up different types of analysis.
 - Able to view aggregated data from the department; including historical trends.
- **Deans**
 - Able to customize a set of items to be added for all faculty at the college level.
 - View same core information as the faculty member and chair.
 - Able to view aggregated data from the college for across department comparisons.
- **Provost**
 - System chosen from faculty-driven committee based on research as well as student, faculty and administrative feedback.
 - Set of simple yet reliable core items consistent across the university.
 - Simple integration with RIT systems.
 - Cost-effective method of implementing university-wide ratings system.

Recommendation #2: Use the same set of established core items across the university that were used in the pilot ($\alpha = .93$ from pilot).

One self-evaluative item, for the purpose of gauging the observational capacity of the student:

I attended this class regularly.

Seven core items:

1. The instructor enhanced my interest in this subject.
 2. The instructor presented the course material in an organized manner.
 3. The instructor communicated the course material clearly.
 4. The instructor established a positive learning environment.
 5. The instructor provided helpful feedback about my work in this course.
 6. The instructor supported my progress towards achieving the course objectives.
 7. Overall this instructor was an effective teacher.
- (Scale: 1=Strongly Disagree to 5=Strongly Agree)

Two open-ended questions:

1. What did this instructor do well?
2. How can this instructor improve?

Bank of additional formative items:

Customizable items may be added and adapted by the college, department and/or faculty member, with the best practice recommendation that the time to complete the evaluation does not exceed 10 minutes for a typical student.

These additional items may be newly created or selected from an item bank. A database of items is available online to assist instructors with item choice.

Rationale for these choices:

The self-evaluative and core items chosen for the SmartEval student ratings of teaching effectiveness pilot study of Fall quarter 2012 were formulated during the Spring quarter of 2012, and are based on (1) feedback from the *AAC Clipboard Survey: University Wide Student Evaluation of Course and Instructors, April 2011*, subsequent to the original 2010 Academic Affairs Task Force recommendations, (2) the email discussion initiated by President Destler in April 2011 as a response to the Clipboard survey, (3) responses from the RIT Colleges regarding the 2010 Taskforce Report, (4) *Student Ratings of Teaching: A Summary of Research and Literature (IDEA Paper #50)* by S. L. Benton and W. E. Cashin, (5) *Student Ratings of Teaching Effectiveness: Using the OIRA Item Bank to Create Your Own Form*, from the Office of Institutional Research and Assessment, Syracuse University <http://oira.syr.edu>, and (6) *Reflective Faculty Evaluation: Enhancing Teaching and Determining Faculty Effectiveness* by J. A. Centra (1993). A detailed discussion of the rationale for changing the self-evaluative and core items from the original 2010 study is given in Appendix A.

Recommendation #3: Provide the online results for an individual instructor (except for instructor added items) only to the instructor, instructor's immediate supervisor and dean, the provost, and tenure and promotion committees per college guidelines.

Rationale: The committee recommends that access to online results be limited to the instructor, instructor's immediate supervisor and dean, and the provost. This is in line with current policies and practices. It is also in line with commonly administered systems in higher education, where approximately 82% of institutions do not share results beyond the faculty member and/or his/her supervisors (department chair, dean and provost). Access to

student rating information required for mid tenure, tenure and promotion committee review, per Policies E.5 and E.6, should also be made available per each college's tenure and promotion guidelines.

Recommendation #4: Re-evaluate recommendations 1-3 after three years of data collection with SmartEvals.

Rationale: As this is a new online system with a new set of core items the committee recommends a re-evaluation of recommendations 1-3 after a 3-year trial period. The three years will allow for adequate data collection with SmartEvals to examine the core items to confirm the reliability and validity of the core items and to evaluate the extent to which SmartEvals is fulfilling the expectations of the RIT community for an online system of gathering student ratings of teaching effectiveness.

V. SURVEY ADMINISTRATION

A. Personnel

- Identify and train campus coordinator/liaison, with vendor (for responsibilities see Appendix C).
- Identify which university and college offices will be assigned to support administration of the student rating system and how they will collaborate.

B. Procedures

- Faculty, departments and colleges will have access to surveys to add items from weeks 1-7.
- Surveys will be open for responses weeks 13-15.
- Faculty will receive access to reports the day after final grading is due for the term.
- Supervisors will receive access to reports one week after instructors.

VI. ESTABLISHMENT OF AN EFFECTIVE SYSTEM OF STUDENT RATINGS

In the course of the committee's work, it became apparent that there are other actions needed to ensure that the student rating system is supported and effective. A companion supplemental report is attached that outlines a number of best practices and fundamental strategies that deserve consideration and implementation.

References

Benton, S. L. & Cashin, W. E. (2009). Student Ratings of Teaching: A Summary of Research and Literature (IDEA Paper #50). Retrieved from http://www.theideacenter.org/sites/default/files/Idea-Paper_50.pdf

Centra, J. A. (1993). *Reflective Faculty Evaluation: Enhancing Teaching and Determining Faculty Effectiveness*. Jossey-Bass, Inc.: San Francisco, CA.

Student Ratings of Teaching Effectiveness: Using the OIRA Item Bank to Create Your Own Form (n.d.). Retrieved March 5, 2013 from the Office of Institutional Research and Assessment, Syracuse University <http://oira.syr.edu>

Appendix A

Original 2010 self-evaluative item:

I had a strong commitment to this course

Original 2010 core items:

1. The instructor motivated me to learn.
2. The instructor was organized and prepared.
3. The instructor communicated clearly.
4. The instructor effectively demonstrated knowledge related to this course.
5. The instructor evaluated my work in a fair and useful manner.
6. The instructor cared about my learning.
7. The instructor was very effective. (design a different scale for final item)

Scale:

- 5 Strongly Agree
- 4 Agree
- 3 Neutral
- 2 Disagree
- 1 Strongly Disagree
- NA= Not Applicable

Rationale for changing the self-evaluative and core items:

It is important to refer to Centra's (1993) seminal research which informs the commonly asked question about what aspects of teaching a student can reliably assess. Centra's work points out that there are six major categories of student evaluative items:

1. Course organization and planning
2. Clarity and communication skills
3. Teacher-student interaction
4. Course difficulty and workload
5. Grading and exams
6. Student self-rated learning

The student **self-evaluative item** (*I had a strong commitment to this course*) as given in the original 2010 recommendation could be indicative of a large variety of interpretations and is thus not a good measure of self-reporting bias. For example, the student may have a strong commitment to the course solely because it is required for the major, but may otherwise have no interest. On the other hand, a student might take a course due to a strong interest in the subject matter without having a strong commitment to it. In addition, commitment is not necessarily tied to ability to evaluate – this would be especially the case for a General Education required course. The replacement item (*I attended this class regularly*) is a more

objective measure of the student's credibility regarding evaluation because it is related to a quantifiable measure of observation.

Item 1 (*The instructor motivated me to learn*) falls into Centra's category 6: Student self-rated learning. Many of the Clipboard respondents noted that it is the student who must supply the motivation, not the instructor. Also, sometimes motivation can be provided in a negative way that may not be beneficial to the student. The original item does not take into account how cultural differences among students may provide for a variety of interpretations. The replacement item (*The instructor enhanced my interest in this subject*) relates to the subjective experience of the student, as provided by the instructor.

Item 2 (*The instructor was organized and prepared*) falls into Centra's category 1: Course organization and planning. The original item is "double-barreled" as noted by several Clipboard respondents as well as by several RIT College responses. It is generally considered poor survey practice to include more than one measure in a single item. It is possible that an instructor was organized but not prepared, or prepared but not organized. This makes a meaningful analysis difficult, if not impossible, because we don't know what the student responded to. The replacement item (*The instructor presented the course material in an organized manner*) considers "organized" to be the more meaningful measure, because "organized" usually involves preparation. In addition, students cannot objectively evaluate an instructor's level of preparation – sometimes an extraordinary amount of preparation on the instructor's part appears to be very impromptu from the point of view of the student.

Item 3 (*The instructor communicated clearly*) falls into Centra's category 2: Clarity and communication skills. This item poses a serious interpretation problem, particularly for NTID students, as noted in the response from NTID. The item could also negatively bias against international instructors who may present the material clearly, but in a manner other than verbal. The replacement item (*The instructor communicated the course material clearly*) focuses more on the presentation of material and leaves open the possibility of other effective modes of communication, depending on the needs and learning modalities of the students.

Item 4 (*The instructor effectively demonstrated knowledge related to this course*) does not neatly fall into any of Centra's categories, but perhaps relates most closely to category 3: Teacher-student interaction. The Clipboard respondents had many negative comments about this item. The majority of objections stemmed from the feeling that students are not qualified to objectively evaluate the quantity and quality of the instructor's knowledge of the subject matter. The implication of this item – that students do indeed possess this ability – was deeply offensive to many faculty. The replacement item (*The instructor established a positive learning environment*) takes into account the student's experience related to the classroom environment as a whole, and the sensitivity of the instructor towards the educational needs of the student.

Item 5 (*The instructor evaluated my work in a fair and useful manner*) falls into Centra's category 5: Grading and exams. Again, a double-barreled item: fair but not useful? Useful but not fair? There is no way to understand the intention of the student's response or interpret it. Also, "fair" is a loaded term for students – literature shows that students are typically not

effective in evaluating fairness – they tend to relate fairness to grade expectations; i.e., if they expect a good grade then the instructor must be fair, but if they expect a poor grade then the instructor must be unfair. The replacement item (***The instructor provided helpful feedback about my work in this course***) focuses more on the helpfulness of the evaluative process. This includes whatever contributed to the student’s learning experience – promptness, relevance, coverage, etc. Overall, did the student perceive the evaluation of their work to be helpful to them personally? If not, the student could expand on this in the open-ended question.

Item 6 (*The instructor cared about my learning*) falls into Centra’s category 3: Teacher-student interaction. Another highly objectionable item, due mainly to the implication (as noted by the Clipboard survey respondents) that the student can somehow magically know what the instructor is thinking or feeling about any individual student or the class as a whole. Also, many Clipboard respondents noted that it is the students who should care about their own learning, at least as much as the instructor does. The original item seems to reinforce the idea that students have little or no responsibility toward their own education – it is all on the shoulders of their instructor. The replacement item (***The instructor supported my progress towards achieving the course objectives***) emphasizes that the student has an ability to observe the instructor and surmise from the teacher-student interaction if he or she supported progress towards the stated course objectives, but cannot objectively judge the instructor’s internal state.

Item 7 (*The instructor was very effective*) falls into Centra’s category 6: Student self-rated learning. This is an overall evaluative type of item, and may be unnecessary, given the responses to the previous items. Does it really provide any additional useful information about the effectiveness of the instructor? Also, many Clipboard respondents claimed that “very” implies a different scale – if the instructor was only effective (what does that mean?) then they would probably score low or mid-range on the Likert scale if students try to calibrate their responses. What really is the difference between “effective” and “very effective?” In addition, the item is too broad to be useful – effective at what? Students might interpret an “effective” instructor as one who enabled them to get an ‘A’ in the class, regardless of what they have learned. Many times students will not be able to objectively evaluate an instructor’s effectiveness until many years after they have graduated, or at least until they are upper-classmen. The committee recognizes this difficulty, but after much discussion, decided that it was important to maintain Centra’s self-rated learning criteria as a survey item, at least as an overall measure of effectiveness from the student’s current point of view. Therefore, the replacement item (***Overall this instructor was an effective teacher***) maintains the summary evaluation from the original statement, but eliminates the vague and un-calibrated “very” qualifier.

Appendix B

Student Pilot Survey for feedback on SmartEvals and IDEA Center

Strongly Disagree Disagree Neutral/Agree Strongly Agree

1. I understood the objectives of the survey.
2. I felt comfortable expressing my opinion by responding to the survey items.
3. The wording of the survey items was clear.
4. The format of the survey was confusing.
5. The survey items appeared biased.
6. The survey was too long.
7. The survey was easy to fill out.

Open – Ended

8. What topics/questions related to teaching effectiveness were overlooked with this survey, if any?
9. What additional comments do you have about this online system of student feedback on teaching effectiveness?
10. What factors influence whether you respond to evaluations of teaching effectiveness?

Faculty Pilot Survey for feedback on SmartEvals and IDEA Center

Strongly Disagree Disagree Neutral/Agree Strongly Agree

1. The survey was easy for me to set up.
2. Overall, the pre-established items reflected aspects of teaching that are important in formulating ratings of effectiveness.
3. The items were sufficient to provide me with valuable information to enhance my teaching effectiveness.
4. I understood how to interpret my ratings report.
5. This student ratings report provided me with useful feedback to enhance my teaching effectiveness.
6. The vendor's resources would assist me in following through on suggestions for enhancing teaching effectiveness.
7. The vendor used an appropriate number of emails to provide me with information.
8. I added items to the pre-established survey. No Yes
9. I filled out my faculty information form. No Yes

Based on all the knowledge you have gained using both surveys and reading reports from both product vendors, please answer the following question.

10. RIT must choose a student ratings product. Which would you recommend?

SmartEvals IDEA Center

Open-Ended

11. Please let us know why you have chosen this product.

Appendix C

Responsibilities as listed by SmartEvals

We have found it to be incredibly beneficial to have one primary contact at each school, and one key person in the IT staff. This simplifies communication between the school and our company, and lets everyone at your school know who to contact with questions. The SmartEvals team is always available to provide support and training for the Primary Evaluation Administrator and the IT contact person, to ensure that everyone at your school is completely thrilled with the new system. Different customers break up these responsibilities differently, so please feel free to call me if you want to discuss how this will apply at your school. Some clients operate with only one point person, and it works perfectly fine. If you have any questions at all, please do not hesitate to call our office.

The SmartEvals.com Team
(888) 309-6373, or (716) 801-1111

PRIMARY EVALUATION ADMINISTRATOR

Qualifications

- Person should be able to access registration information at the school
- Person should be comfortable with computers and using the internet frequently

General responsibilities

- Act as the primary contact person at the school for the evaluations
- Communicate as needed with the SmartEvals staff
- Be available to teachers, administrators, and students to handle evaluation questions
- Coordinate activities to ensure a high response rate, based on SmartEvals' suggestions

Before the first evaluation period

- Gather copies of any and all question sets (evaluation forms) that will be used
- Input administrators and unit heads who need access to certain areas of the data
- Communicate with appropriate individuals at the school to set login procedures, evaluation dates, information release date, and other settings.
- Receive information from SmartEvals that need to be distributed to faculty/administrators, and distribute them via email and/or campus mail
- Train with SmartEvals staff on the administration pages of our system, to prepare for the evaluation period.
- Follow our step by step guide to generating outstanding response rates.

During each evaluation period

- Be available to answer basic questions from users. Ex: "help, I forgot my password." (This may be a responsibility of the IT Contact Person instead, or delegated to a general help desk.)
- Compose reminder messages to students and teachers, and send them through the SmartEvals system
- Track response rates and communicate with professors

Monitor response rates and be in communication with the SmartEvals team

I.T. CONTACT PERSON

Qualifications

- Person should be able to access registration information at the school, and be comfortable writing queries / exporting data into a flat text file

General responsibilities

- Create the upload file each semester with correct registration information
- Communicate as needed with the SmartEvals staff

Before the first evaluation period

- Communicate as needed with the SmartEvals staff, to learn our requirements for the upload
- Work with the registration system to prepare the first upload file
- Ensure that all users (students, teachers, and admins) have a login ID and password for the initial upload
- Upload the file to the SmartEvals system
- Work with SmartEvals staff members to set up single sign-on through the campus portal

During each evaluation period

- (This may be the IT person, OR the Primary Evaluation Administrator)
Be available to answer basic questions from users (Ex: “help, I forgot my password”)

Again, our job is to work with you to delegate these responsibilities to the most appropriate individuals. We are happy to create different breakdowns to fit your school, but we have found these job responsibility assignments to be successful.

Appendix D

STUDENT RATINGS PRICE QUOTE SMART EVALS (Ronald Jennings – Smart Evals.com)

Please find the price for a 3-year service agreement below. Years 2 and 3 reflect our standard 3.5% annual cost-of-business increase:

Year-1	\$19,995
Year-2	\$20,695
Year-3	\$21,419

The price I provided does include everything. I've provided a list of specifics below:

- Setup of the service.*
- Training of your Program Administrators, Faculty, and Unit heads, Administrators and Instructors.*
- Subscription to the SmartEvals.com online course evaluation message board.*
- Hosting the service on our secure servers.*
- Follow-up question technology.*
- Use of the Dropped Course survey.*
- Setup for LDAP authentication or Single Sign-on authentication through student portal.*
- Customizable reporting.*
- Use of the ASCEND question set.*
- Faculty development tools, specifically the Percentile Rank analysis and the myFocus tool.*
- Email effectiveness analysis.*
- On-going customer and technical support Monday - Friday 9:00 am to 5:00 pm (extended hours during the evaluation period).*
- Response Rate coaching and best practice tips.*