FRWC guidelines for using and interpreting IDEA feedback data in the faculty review process

IDEA course evaluation data will be included in faculty's evaluation for the first time in AY 2015-2016. How these data will be interpreted by review bodies at FDU may differ and thus it is critical that general guidelines be developed to retain some consistency of decision making. These guidelines put forth by the **Faculty Rights and Welfare Committee** of the Faculty Senate are intended to balance the rights of the faculty member being reviewed and those of the relevant review bodies.

We were advised that IDEA was chosen because it provides evaluative data on teaching from the students' perspective, as well as ample developmental data, and data that can be used for departmental or programmatic curriculum development. For the purpose of this document, we are not interested in the developmental or the programmatic data. These data should absolutely be used by faculty and appropriate decision makers, but it is beyond the scope of what should be included in a faculty portfolio, and what is used in making evaluative decisions.

The university should make an effort to adapt the IDEA system to the FDU environment. In particular, there should be an option of comparison to courses within the same FDU college or department or even a group of courses within a department such as multiple sections of the same course.

1. IDEA should be used in conjunction with peer and chair observations, samples of student work, evaluation of teaching materials, syllabi and course grade distributions and any other artifacts that may be indicative of teaching effectiveness. As IDEA states their data should constitute no more than 30%-40% of what evaluative bodies consider when evaluating teaching.

2. Given the inherent unreliability of small sample sizes, decisions should be based on consistency of scores. An occasional low score should not be over-interpreted.. Context is important, and needs to be considered. Finally, small differences, say between a 3.9 and a 4.1 should not be over-interpreted, considering the noise inherent in any student survey.

3. The faculty should have a choice to use the raw scores or the adjusted scores (or t-scores) to make his/her case.

4. Faculty with an adjusted score below a certain level should be required to provide some explanation and/or plan for improvement. The departments should be given some freedom to determine what that trigger level should be, depending on the type of course in question. For some departments a t-score of 40 may be a proper trigger level. However, if the adjusted scores in a given discipline are typically below the mean then a lower trigger level may be more appropriate. Once an option of comparison to groups of FDU courses is available the trigger level will be easier to define.

Appendix

Here's the information from the IDEA reports that faculty will be presenting in their portfolios:

First Page:
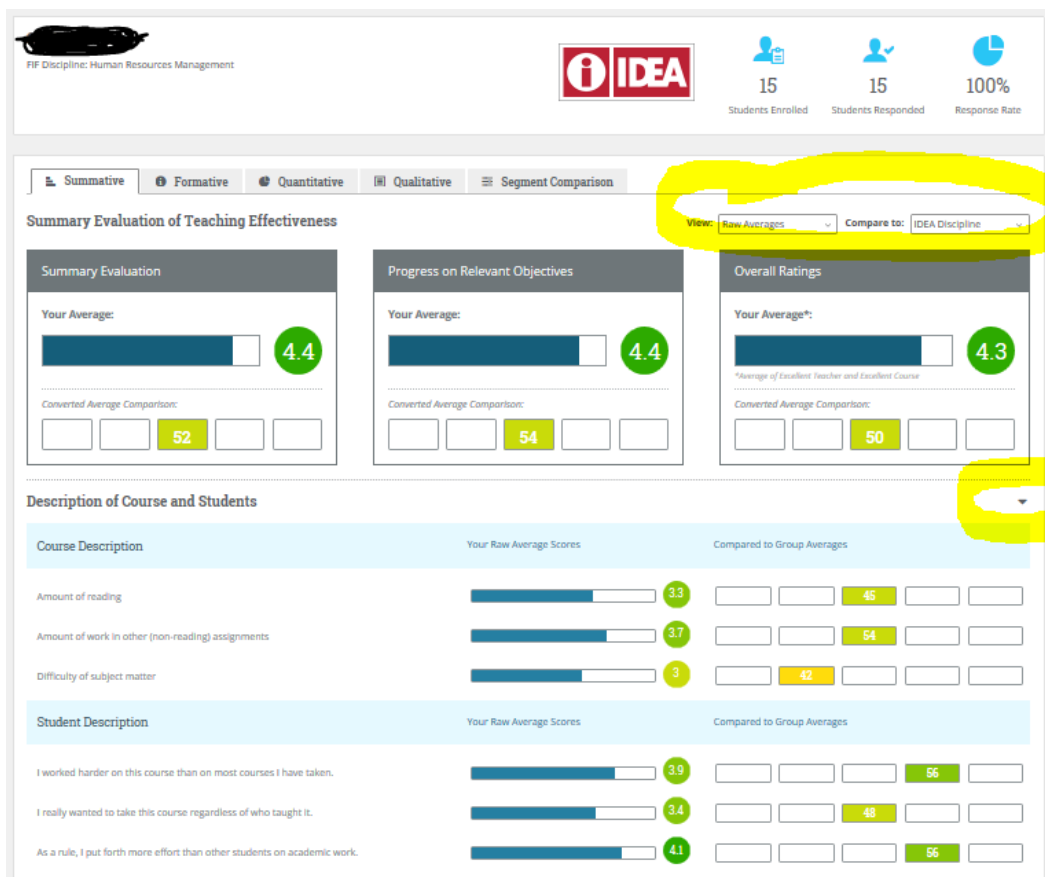
- All numbers in circles are ratings on a 5-point scale
- All numbers in rectangles are t-scores, set so the mean is 50 and the standard deviation is 10. (This means a score of 40 is one standard deviation below the mean, or a score of 55 is ½ of a standard deviation above the mean).

First, they will show the main summary page. Here, you see the main "headline" numbers.

- This page should show "raw averages", meaning there are no "adjustments" made to the numbers based on class context.
- And it should show "comparison to the IDEA discipline," so responses are compared to the population of courses in the same discipline throughout the 400+ colleges that use IDEA.

Progress on Relevant Objectives is based on the students' self-assessment of how much progress they made only on the learning objectives the faculty member chose as important

or essential for the course. This is somewhat analogous to "Pedagogical Skill" under Endeavor.

Overall Ratings is based on a couple of summary items, and is somewhat analogous to "Rapport" from Endeavor.

Summary Evaluation averages these two categories together into a single score.

Description of Course and Students- The items and scores underneath are the contextual factors that IDEA uses to "adjust" scores up or down based on factors outside the instructor's control.

> *In general, small classes of bright, highly motivated students who actively chose the class, and who found the work relatively easy will give HIGHER ratings independent of how well the professor teaches; large classes with dull, unmotivated students, forced into a class they don't have an interest in and find that class difficult with a lot of work will give LOWER ratings, independent of how the professor teaches.*

Again, we are first looking at Raw averages- that do not adjust scores based on these factors. However, it is useful to see if there is anything remarkable about the context of the course. This way, PRCs can be made aware of extenuating circumstances or issues that may have affected ratings. Faculty may also provide a narrative about the context or other potential influences on scores. For example, this may have been a brand new class, or 40% of the students were caught plagiarizing and took revenge during evaluations.

If a faculty member believes that the adjusted scores are a better reflection of their teaching effectiveness, they will also submit the "**First Page**" information using **Adjusted Averages.** Here's what these look like for the same class as above:
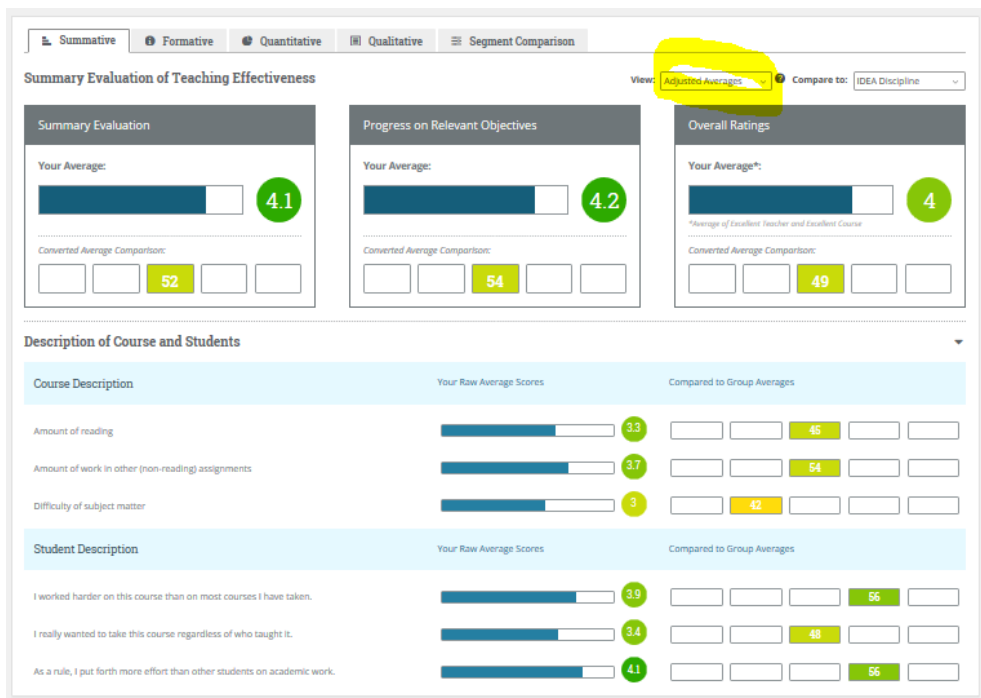
In this case, the course context was relatively favorable, so the adjusted averages are very slightly lower than the raw averages. Like this, in most cases, adjusted scores do not differ very much from raw scores. There, of course, may be situations in which context does have a stronger effect. Please be mindful of course context.

Second Page:
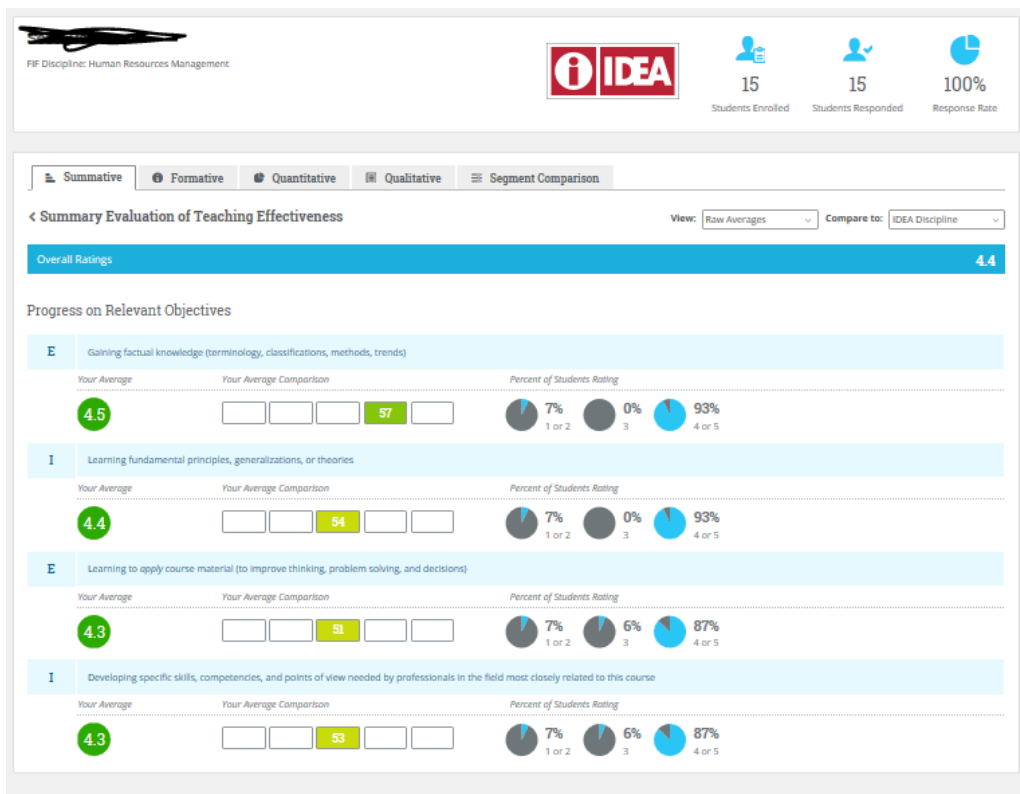
Next, the faculty will present their scores indicating student progress on the Learning Objectives they selected as **Important** or **Essential** to the course. Most faculty should have selected around 3-5 learning objectives but they may have chosen more or fewer.

Again, scores in circles are on a 5-point scale, and scores in rectangles are t-scores. You can also see the distribution of scores in the three pie charts to the right.

This is the heart of IDEA evaluations. **Regardless of how a faculty member teaches, are they doing what is necessary to help students make progress on the most important learning objectives for their courses?** If these scores are good, you can have confidence in this instructor. If they are lower than you'd like, now you have some information to provide guidance, for example, "*In next year's review, we'd like to see improvement on X*."

If all t-scores are above the set trigger level, this is all a faculty member is required to report, and this information should be sufficient for a personnel review body to use for decision-making. Faculty may, of course, submit additional information, and you may request additional information as you see fit.

However, if **any t-scores on this or the previous page are below the set trigger level**, the faculty member should provide an explanation and a plan for improving on those scores. They can also choose to provide a 3rd set of data for review (see the next page).

Third Page (if needed)

If any t-scores are below the set trigger level, the faculty members may also choose to provide one more set of scores. They may provide the **Segment Comparison** data for that course, in which they provide information on that course as compared to all **courses within their discipline in their department**. It will look something like this.
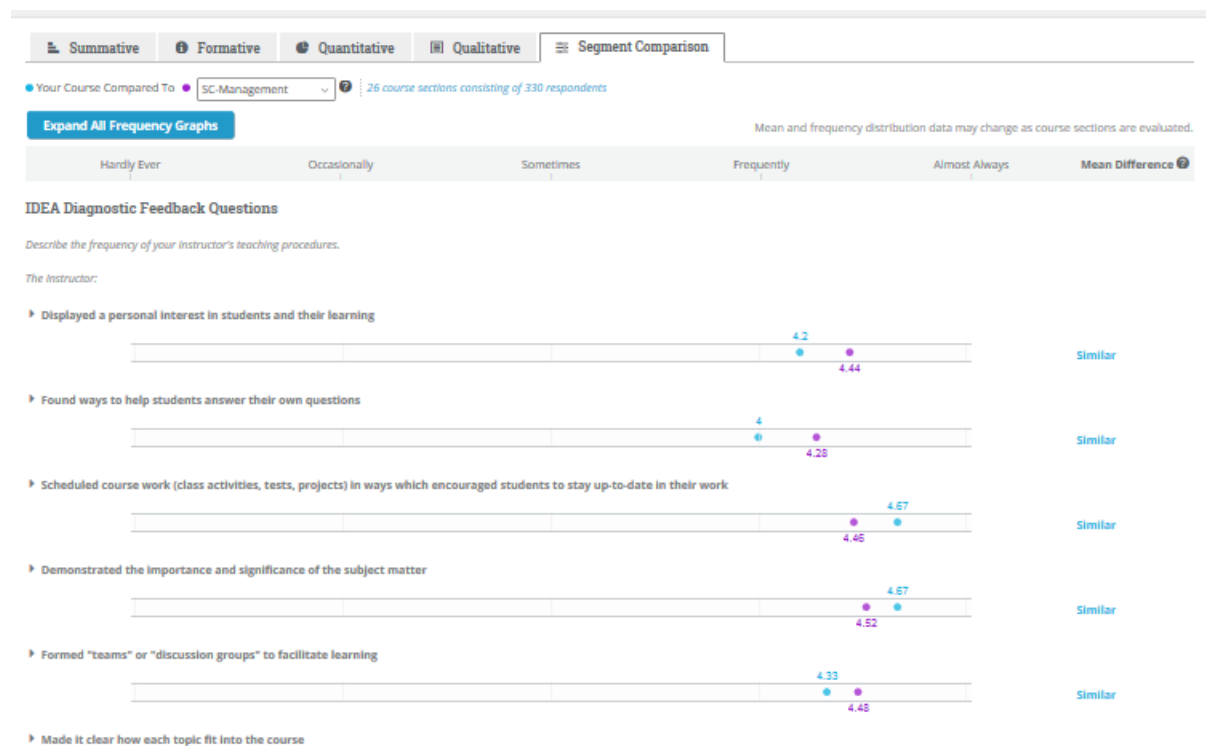
Please note that this data includes all items on the IDEA survey, including learning objectives NOT SELECTED as important or essential, and teaching behaviors that MAY or MAY NOT be relevant to that particular course. Please use discretion when making evaluations based on these items. You will also get data on all contextual questions- this report will be several pages long.

The dots that show the mean for this instructor/class are plotted alongside that of the average of all other courses of the same FDU discipline within department (*for example, in this case, this management class is compared to all other management classes within the department of management- which means it is not compared to exec or phar classes within management. All departments will differ, but faculty have been instructed on how to get you the correct information*.) Finally, the right hand side shows whether the instructor is significantly different from the rest of her/his peers. This additional information will be useful for those cases in which additional scrutiny is required based on relatively low "headline" numbers.

Faculty will also provide a narrative explanation and plan for improvement if their headline t-scores are below the set trigger level.

Endeavor Scores

Finally, all non-tenured, tenure track faculty have had the Endeavor course evaluation items added to their student surveys. Faculty used an excel spreadsheet we have provided them in order to derive **Rapport** and **Pedagogical Skill** scores and ensure consistency between how they were evaluated before and after the changeover to IDEA. Once a faculty member is tenured, they will no longer have Endeavor items added to their surveys, unless they specifically request to do so.

They will provide the result of their spreadsheet, which should look something like this:

Endeavor Questions

1. Never	2	3. Seldom	4. Sometimes	5. Often	6	7. Always	Did Not Answer	Total Responses	Average
The student had to work hard in this course.									
1	2	3	4	5	6	7		28	5
Each class period was carefully planned in advance.									
3	4	1	5	8	4	3		28	4.25
Class discussion was welcome in this course.									
0	1	5	5	5	7	5		28	4.96429
The student was able to get personal help in this course.									
0	2	0	0	6	10	10		28	5.85714
The instructor presented the material clearly and summarized major points.									
0	2	1	0	5	8	12		28	5.85714
The grading accurately reflected the student's performance.									
0	0	0	7	7	7	7		28	5.5
The instructor answers questions well.									
0	0	0	7	3	5	13		28	5.85714
Exam questions reflect the content of the course.									
0	0	0	1	3	4	20		28	6.53571
Classes meet regularly and for the whole period.									
0	0	0	3	4	9	12		28	6.07143
The course materials (text, readings, equipment …) were helpful.									
0	0	0	4	3	3	18		28	6.25
This course has increased my knowledge and competence in this area.									
2	3	0	0	0	8	15		28	5.75

How many hours do you spend on this course outside the classroom per week?

0 hours per week	1 hour per week	2 hours per week	3 hours per week	4 hours per week	5 hours per week	6 hours per week	7 hours per week	8 hours per week	9 hours per week	Did Not Answer	Total Responses
											0

1	hard worl	5	Rapport	5.6607	
2	advance	4.25	Pedagog	5.2643	
3	class dis	4.96429			
4	personal	5.85714			
5	presenta	5.85714			
6	grade ac	5.5			
11	increase	5.75			

PRCs should interpret these scores as they always have. Faculty, of course, may also provide an accompanying narrative, but they do not have to.

Taking into consideration our large international student body, we agree that the IDEA, language utilized – for questions posed about "instructor"- may be too complex for students with English as a second, third or fourth language to comprehend. We suggest for questions to be added to this section with short responses in a simpler English language, i.e. two or three words to describe instructor and/or course.