

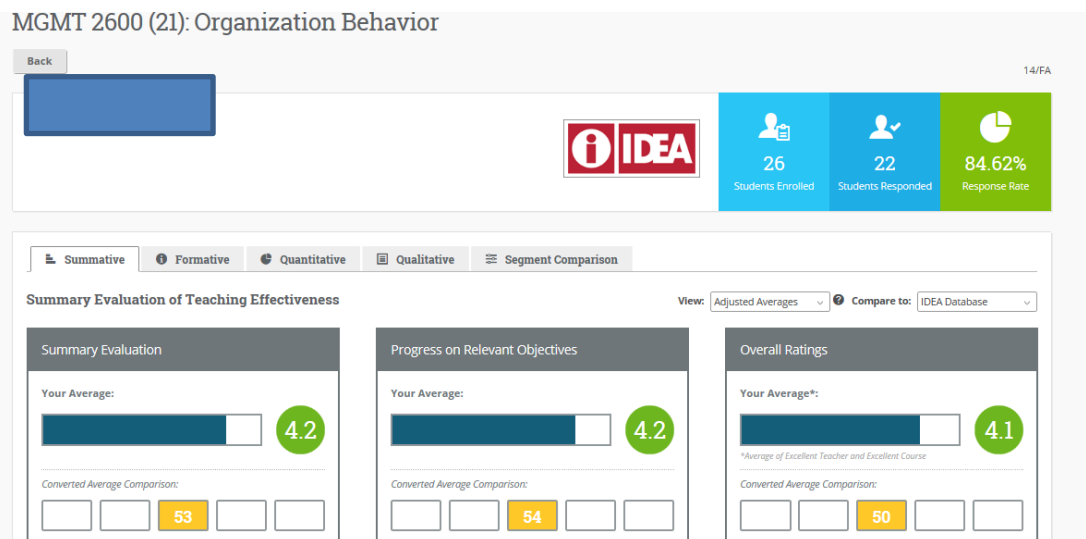
If I Were to Interpret an IDEA Evaluation Report for Evaluation Purposes

First, student evaluation of teaching effectiveness should be NO MORE THAN one-third of how we evaluate a faculty member's teaching. Portfolios, observations, samples of assignments and student work, etc. should all be considered when making a full judgment.

For that no more than 1/3, if I were a department chair, or member of a DPRC or CPRC, this is how I would look at IDEA results in making my evaluation. Each academic unit should develop its own standards according to your particular academic situations. IDEA is simply a different tool, judgment of faculty peers is still a fully academic matter. This is only a recommendation.

Summative Information

First, I would look at summative information:



In this case, the big picture view is that this course was taught well, and the outcomes are largely in line with what a chair, DPRC or CPRC would like to see.

Summary Evaluation: 4.2

This is the closest metric we have to an overall score, although it would be reductive to winnow all the information into a single number. This score combines the progress on learning objectives items (50%) with the overall ratings items (50%).

4.2 on a 5 point scale is, on its face, good. This represents a T-Score of 53. T-scores have the mean at 50 and each standard deviation is plus/minus 10. In this case, the overall score of

4.2 represents +.3 sd above the mean or all classes taught at the 4000+ universities that use IDEA. This is somewhere around the 58th percentile.

Progress on Relevant Objectives: 4.2

This is the most important consideration for me.

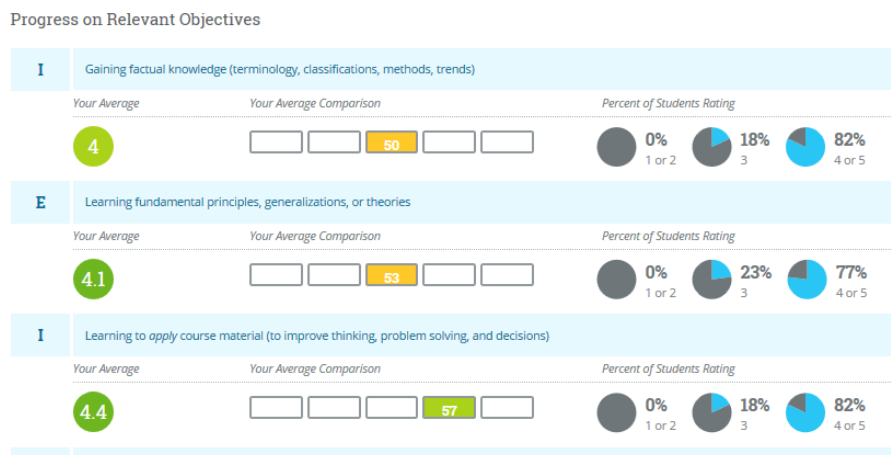
This score indicates that students rated their progress on ONLY the learning objectives that the faculty member identified as Important or essential as 4.2 on a 5 point scale. Again, on its face, this is good. Note the T-score of 54, about the 66th percentile.

In short, this instructor should be rated as effective, and it is likely I would give this person a positive review, if all the other teaching documentation were also in good shape. There is opportunity for development and improvement, but, as this guide is focused on evaluative purposes, I'll focus just on evaluation (We are developing a companion document for developmental feedback).

In fact, a good rule of thumb is that instructors with t-scores listed in the yellow box (T-scores ranging from 45-55) are effective teachers, representing the middle 40% of scores nationwide. T scores in the next box to the right include T Scores of 56-62, which represent the 70th through 90th percentiles. The far right box contains T-scores for the top 10 percentile of classes nationwide. Instructors in these categories should be seen as highly effective.

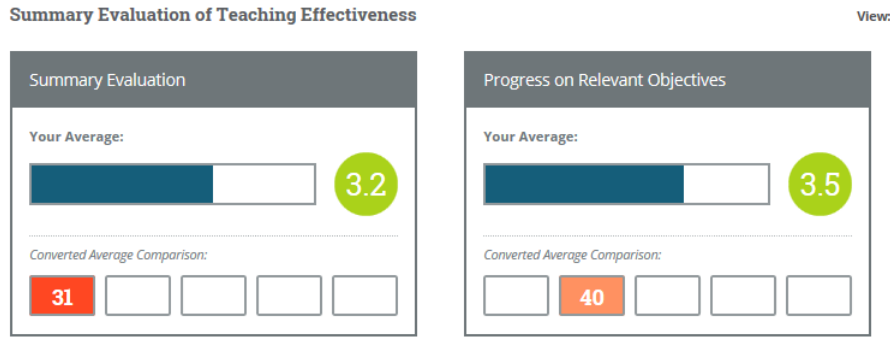
Similarly, the boxes to the left are below nationwide averages. Faculty who consistently earn these scores are not as effective as they could be. Chairs and senior colleagues should work with the faculty member on their development in the classroom, and if there is no improvement over time, this faculty member risks a negative evaluation. Again, context and other teaching materials also matter.

You can click on the Progress on Relevant Objectives box to see this:



In this case, you can see that the students reported good progress on these learning objectives. The 5-point scale score is in the green circle, the t-score for each item is located in those rectangles, and the pie charts show the percentages of students rating 1-2, 3, or 4-5 on each objective. Again, these are good scores, indicating that students felt they made good progress on the most important objectives for the course.

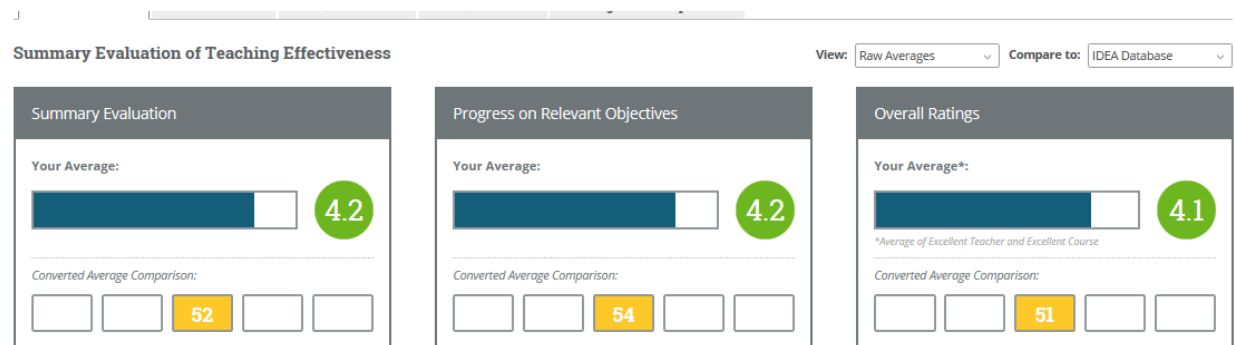
Here's an example of a less effective class:



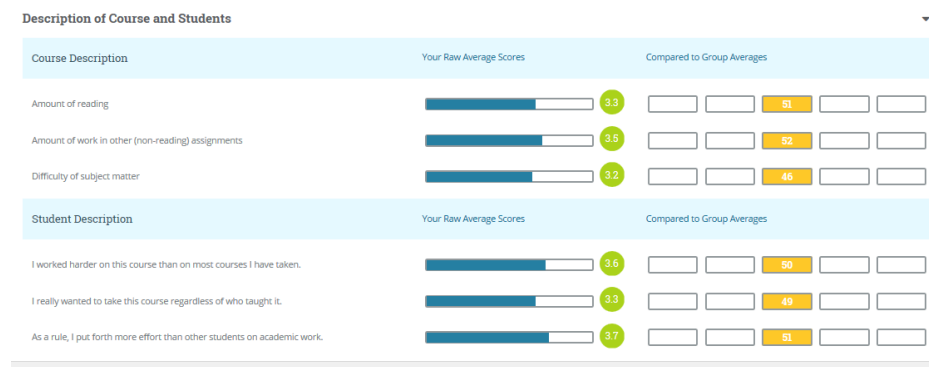
What would you say in this case? First, I would drill down to the particular learning objectives, as noted earlier. Then, I'd want more information about the class and whether these results are typical. If they are, this would indicate a potential problem. A closer analysis of quantitative information and segment comparison data is needed here, along with examination of class materials and perhaps an oral or written explanation from the faculty member.

Raw vs. Adjusted Data

So far, we've looked at adjusted data, after adjustments for classroom context beyond the instructor's control such as class size and student study habits. Excellent students often report more progress on learning objectives than marginal ones, and those in small classes rate teachers more highly than in large lecture halls, for example.



We should do a quick check to be sure the raw scores are also ok. In this case, the raw scores are almost exactly the same as the adjusted ones. In fact, it is unlikely that the raw and adjusted scores will differ by more than plus/minus 0.3 on the 5 point scale. However, if they do differ substantially, I would need to dig deeper into the context of the course- what was different or idiosyncratic about this course that made the raw and adjusted scores so different. A quick way to see the most likely reasons why raw and adjusted scores may differ is clicking on the small triangle just below the summary evaluation. Doing so opens up this information:



In this case, the characteristics that could lead to score adjustments are all almost exactly average. This class is of medium difficulty, with an average workload and with average student characteristics.

Again, in the vast majority of cases, the differences will be slight. But those who teach small classes of highly motivated students who actively chose to take that class with that instructor do have a context that is conducive to high ratings. Those who teach large classes with marginal students who have no choice but to take this class that they are not interested in have a context that is less conducive to high ratings. The adjustments are an attempt to level the playing field.

In future semesters, we will be able to get comparative data here on the front page to all FDU classes, all classes nationwide within the class' discipline (through the use of CIP codes), and all FDU classes within discipline. Unfortunately, we could not get this data for the first semester.

Formative/Developmental Information

Because everything looks good here on the first page, I have no major concerns about student ratings. However, I would also like to investigate a bit further. The formative information is really helpful for development, and we should take just a quick look there next.

Teaching Methods and Styles

Stimulating Student Interest	Suggested Action
Demonstrated the importance and significance of the subject matter	Strength to retain ▲
Stimulated students to intellectual effort beyond that required by most courses	Strength to retain ▲
Introduced stimulating ideas about the subject	Strength to retain ▲
Inspired students to set and achieve goals which really challenged them	Retain current use or consider increasing ▲

This is just a small snippet of the page. In this section, you will see a listing of 20 different teaching behaviors, broken into a few categories. Each will have a color coded message to the right-hand side.



- A green “strength to retain” means that students reported that this professor used this teaching behavior in a way that was appropriate to the selected learning objectives.
- A yellow message means that things are fine, but there may be room for improvement.
- A red “consider increasing use” message indicates this may be an area for improvement. If I were on a DPRC, I would like to see behaviors flagged with red to be improved upon over time by the faculty member when teaching this same course.

Here’s an example with some areas for improvement:

Encouraging Student Involvement	Suggested Action
Related course material to real life situations	Consider increasing use ▲
Involved students in “hands on” projects such as research, case studies, or “real life” activities	Retain current use or consider increasing ▲

If these red areas are relevant to the chosen learning objectives, this is an area for concern, and one should dig deeper.

One can click on the little triangle on the right hand side and get more information about ways to improve that teaching behavior. If you do, you’d see something like this:

Explained the reasons for criticisms of students' academic performance		Retain current use or consider increasing ▼
Your Average  4.1	1 of 2 Relevant Objectives Learning to <i>apply</i> course material (to improve thinking, problem solving, and decisions)	Suggested Action You employed the method with frequency typical of those teaching classes of similar size and level of student motivation. • POD IDEA Note
Students Rating  5% 82% 1 or 2 4 or 5		

This best thing here is the link to the POD IDEA Note, which, when you click it will open a new window with a helpful white paper written by professors of higher education with ideas on how to improve. This is an especially useful tool for development, and as a starting point for faculty conversations with chairs and senior colleagues.

Quantitative Information

Next, the Quantitative Tab gives the distribution of student responses for the items on the survey. These can be examined for trends. However, please note that this section contains information on all learning objectives and teaching behaviors, even ones that were not considered relevant for the course. If I see a low score, it is not necessarily a concern – we should focus only on the selected learning objectives as we did earlier.

Summative Formative **Quantitative** Qualitative Segment Comparison

Describe the frequency of your instructor's teaching procedures.

The Instructor:

Hardly Ever	Occasionally	Sometimes	Frequently	Almost Always	Mean	Standard Deviation	Did Not Answer	Total Responses
Displayed a personal interest in students and their learning								
0%	4.55%	0%	31.82%	63.64%	4.55	0.72	0	22
0	1	0	7	14				
Found ways to help students answer their own questions								
0%	4.55%	4.55%	59.09%	31.82%	4.18	0.72	0	22
0	1	1	13	7				

In the case of a class with lower scores (as seen below), you could examine these items to reveal that the instructor held students to a high standard and expected students to take responsibility for learning, while using a variety of techniques. So, at least there is something to build on, and perhaps a clue as to why the scores were low.

The instructor expected students to take their share of responsibility for learning.								
0%	0%		16.67%	50%	33.33%			
0	0		1	3	2			
The instructor had high achievement standards in this class.								
0%	0%		16.67%	33.33%	50%			
0	0		1	2	3			
The instructor used educational technology (e.g. Internet, e-mail, computer exercises, multi-media presentations, etc.) to promote learning.								
0%	0%		33.33%	50%	16.67%			
0	0		2	3	1			

Further, you may notice this is a class with only seven students, so the results may not be stable. So, even in this case of a low rating, I would want to look for more information, and look across classes and semesters before making a firm decision

Endeavor Questions

The old Endeavor questions are at the very bottom of this section, just for non-tenured tenure-track faculty and those who requested it. Each item is rated on a 7 point scale, but one would need to use this formula (see the excel worksheet posted to <http://fduidea.wordpress.com> to recreate the “rapport” and “pedagogy” factor scores).

In this case, the faculty member is tenured; we do not have this data.

Qualitative Data

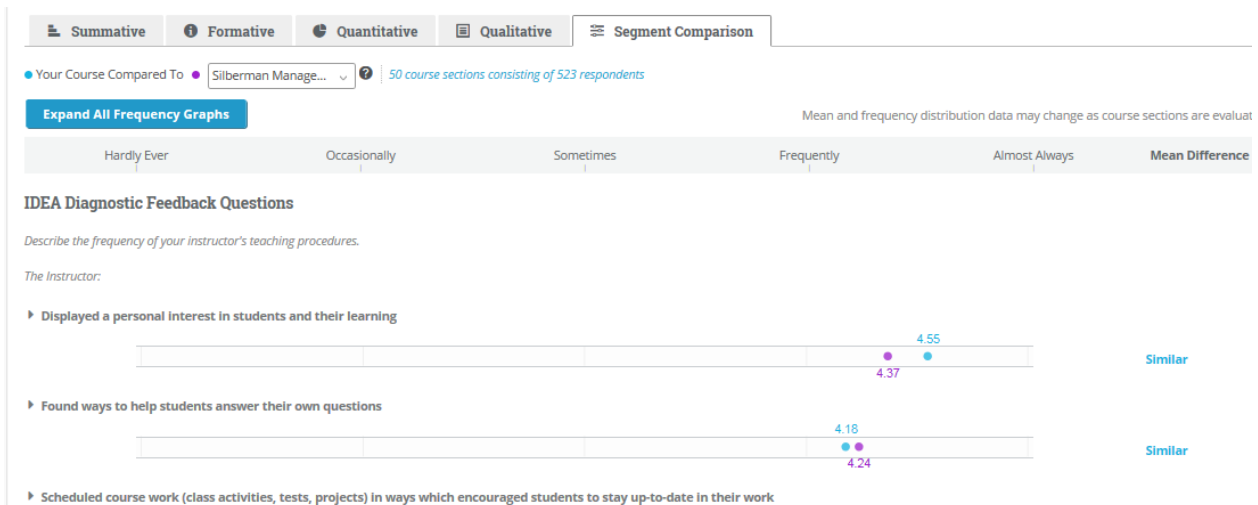
As is consistent with past practice, chairs and PRCs will not have access to qualitative data, unless it is volunteered by the faculty member.

Segment Comparison

Finally, under Segment Comparison, there is a pull-down window so one can look at every question on the survey in comparison to various groups. In this case, we could look at

- Fairleigh Dickinson University
- Office of the Provost (all courses under the direction of University Provost , i.e., all classes at FDU)
- Silberman College of Business (all courses under the direction of SCB Dean)
- Silberman- Management (all courses under the direction of Management department chair)
- SC- Management (as Management also has courses listed as BUSI, EXEC and others, this tab shows only against other MGMT courses)
- All Sections in Course (in this case, there are 6 sections of the course, taught by 4 different faculty members. You can see this section against the average of the other sections here) This may be particularly useful

In this case, we can see this course is receiving responses similar to other SCB courses.



Please note that the only additional data we cannot get JUST THIS FIRST SEMESTER is the comparative data rolled up into the summary scores on the first “Summative” screen. The item-by-item intra-FDU comparison data are all here.

Printing your results

Currently, there is no simple way to print one’s evaluation reports. We are working with IDEA on a technical solution. However, there is a pretty easy work-around we can use in the short-term:

- First you need to download Cute PDF Writer- <http://www.cutepdf.com/Products/CutePDF/writer.asp>
- Then “Cute PDF Writer” becomes one of the printers installed on your computer.
- Then, go to each screen of your IDEA report, and (on Internet explorer) go to the File tab and select Print. Choose Cute PDF Writer as the printer.
- It will convert the entire web page (not just what you see on your screen, but the whole page) to a black and white pdf file. Unfortunately, we lose the color-coding
- You can then print out these pdfs to include them in personnel files and teaching portfolios.

In Conclusion

IDEA reports give much more detailed information about student perceptions of their progress on learning objectives, of faculty teaching behaviors, and of the context of the course.

As this is a new process for us all, we should be sure to evaluate a full amount of information and make sure that student evaluations make up no more than one third of our evaluation of faculty teaching.