

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

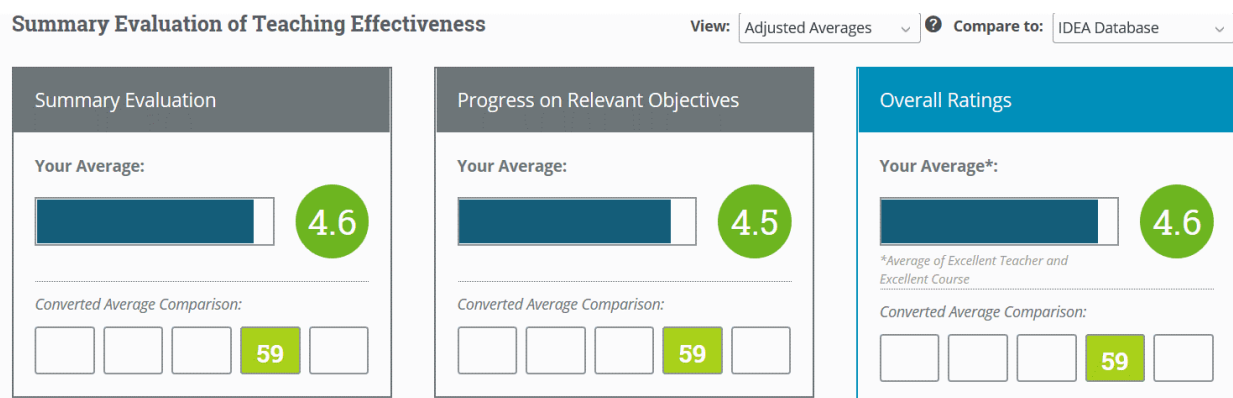
The IDEA Evaluation Report is primarily considered a tool to evaluate course and teaching effectiveness, yet it is also very useful for *developmental* purposes. Once scores are received and areas of strength and those needing improvement are identified, we can use the evaluation reports to aid in better course development and/or teaching success.

It is important to keep in mind that scores do not measure actual course or teaching efficacy, but rather student perceptions of them. Therefore, in cases of poor scores, chairs and faculty members should further analyze score breakdowns in order to reasonably assess which areas should best be focused on for improvement.

In this guide, companion to our earlier piece, “If I Were to Interpret the IDEA Evaluation Report for Evaluation Purposes,” we will systematically discuss how to use IDEA for developmental purposes, considering both faculty member teaching and course effectiveness.

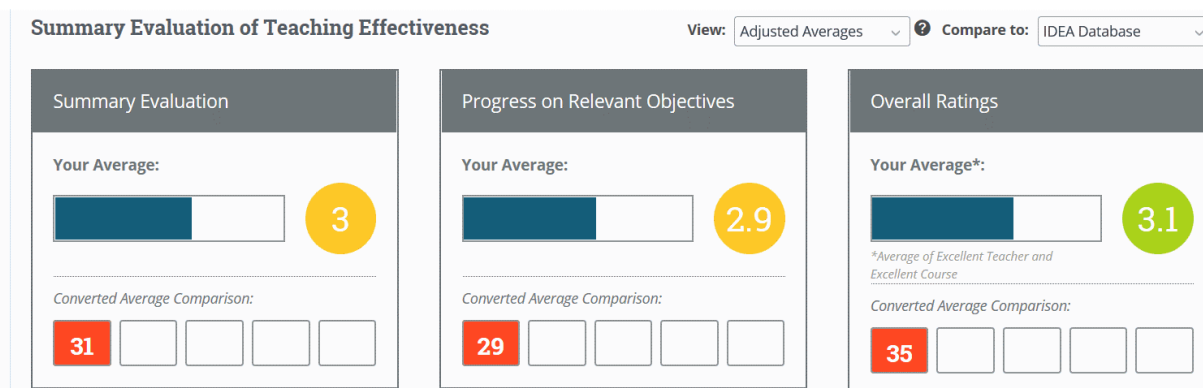
Identifying and Enhancing Teaching Effectiveness

The first scores we’ll tend to look at will be the “Summary Evaluation of Teaching Effectiveness,” scores under the “Summative” tab. Here is an example of what we would consider to be a “good” score:



In most cases, if a faculty member consistently receives above average scores, similar to the example above, we would urge him/her to continue doing what he/she is already doing, and offer congratulations rather than intensive instruction. However, unless the score is a 5, there will always be room for improvement. We’ll soon see where to find resources that will provide helpful guidance.

While faculty members and chairs would be gratified to see the above scores, in some cases they may be presented with scores that look more like this:



What should be done in this case? Our companion guide discusses how to gauge whether the professor should be *evaluated* positively or negatively, however our focus here is on *development* of both the faculty member and the course. Therefore, our first step would be to determine whether the above scores are typical for the particular faculty member. If so, we would move on to the second step of determining why students are generally dissatisfied. It would then be important for chairs and senior colleagues to work with the faculty member on his/her improvement in and out of the classroom, and, if there is no improvement over time, this faculty member would risk a negative evaluation.

The “Formative” Tab

In order to identify areas for improvement, we will begin by clicking on the “Formative” tab.



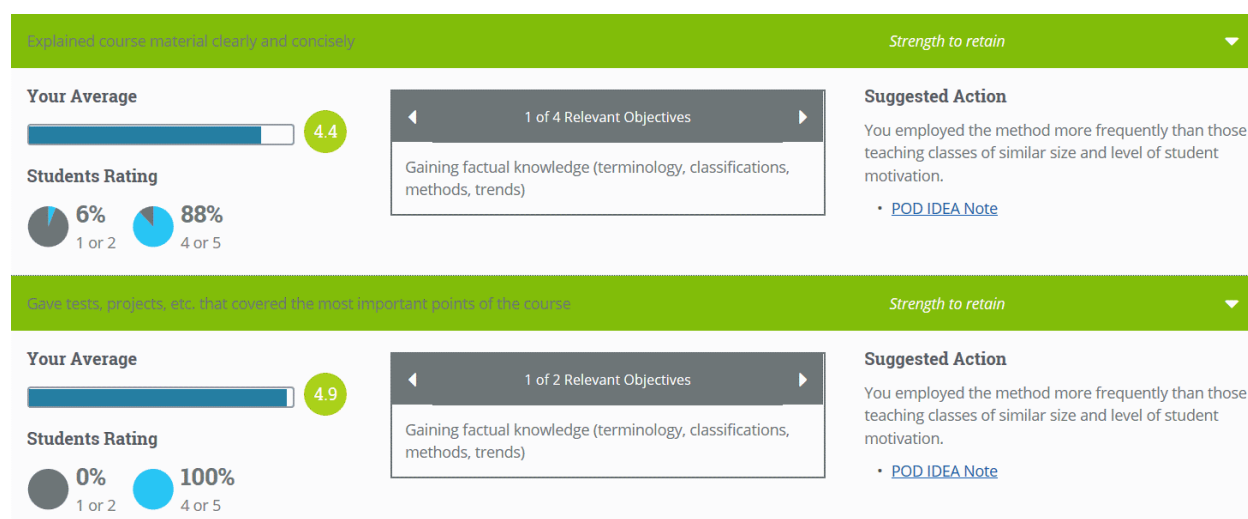
In this tab, the twenty different teaching behaviors, broken into several categories, are enumerated and color coded in order to highlight which goals should be focused upon.

- 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. A Chair or DPRC member would like to see behaviors flagged with red to be improved upon over time by the faculty member when teaching this same course.

Immediately below the color-coded message on the right hand side is a “Suggested Action” which tells the professor, or anyone viewing the report, how frequently effective methods were employed. A blue, [POD IDEA Note \(http://ideaedu.org/research-and-papers/pod-idea-center-notes-instruction\)](http://ideaedu.org/research-and-papers/pod-idea-center-notes-instruction) is provided. “PODS” are written by professors of higher education, who, based on extensive experience and research provide detailed recommendations for how to improve teaching performance. This is an especially useful tool for development, and as a starting point for faculty conversations with chairs and senior colleagues. PODS provide referenced background information, helpful hints, applications for online learning and also address assessment issues.

GREEN TABS

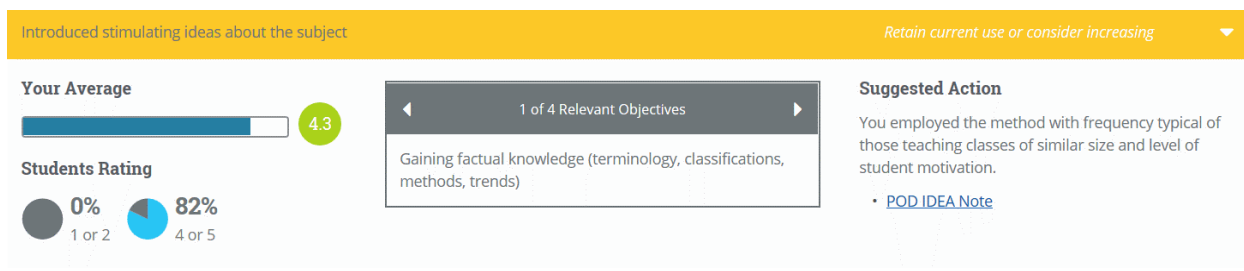
Here is an example of what we’d love to see:



In this case, while the faculty member may find it interesting to explore the [POD IDEA Note](#) for additional insights in order to enhance his/her teaching, the chair and DPRC would probably advise the professor to do more of the same and move on.

YELLOW TABS

While green bars are what we’d aim for, in many cases a yellow bar may not really be much cause for concern either. However, if we did see a yellow bar, we’d want to look a bit more closely at the quantitative information in order to see where improvement is needed most. In the example below, the professor is rated on “introduced stimulating ideas about the topic,” and received a score of 4.3. More importantly, 82% of the class assigned the faculty member a 4 or a 5 on a 1-5 scale. The other 18% assigned a score of 3 – which generally represents a neutral response. Many chairs or DPRC’s would be completely satisfied with that, however, taking a quick look at the [POD IDEA Note](#) may be useful in order to move more students from “neutral” to “satisfied” in the future.

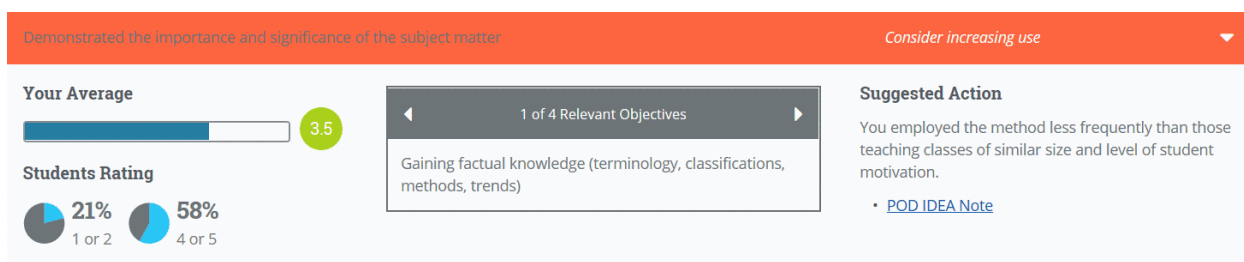
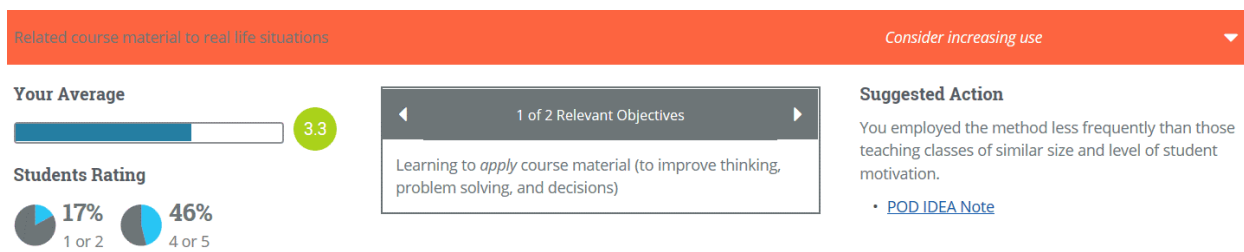


In this case, the [POD IDEA Note](#) painstakingly goes through many different ways for getting students excited about the topic. For example, it recommends:

- 1) Begin with a real-world problem that students can attack and possibly solve...
- 2) Use examples that clearly connect with students' backgrounds
- 3) Stimulate thinking through proper question structuring. Questions asking for opinions, observations, or recall are useful to open a discussion or topic...
- 4) Use humor...
- 5) Use assignments that clearly connect to course content and intended outcomes, and clearly inform students why these connections are important to learning.

RED TABS

Scores that would concern a Chair or DPRC more would look something like the examples below. In the following cases, almost 20% of the class did not feel that the faculty member met the objective satisfactorily. While we all know that there may be several students who will not even read the question but assign the lowest score all across the board, in a case where 20% of the class is unsatisfied, taking steps to improve would be appropriate.



The first of the two red bars here rates “related course information to real life situations.” Were we to click on the [POD IDEA Note](#) provided on the lower right hand side, this is what would pop up:

IDEA ITEM #11: RELATED COURSE MATERIAL TO REAL-LIFE SITUATIONS

POD-IDEA Notes Instruction | Series Editors: Michael Theall, Youngstown State University; Derek Bruff, Vanderbilt University; Amy Gross, The IDEA Center | Author: Michael Theall, Youngstown State University

- Background
- Helpful Hints
- Applications for Online Learning
- Assessment Issues
- References and Resources



Background

There are many reasons for incorporating real-life situations into instruction. Foremost are that *applications of theoretical material in real-life situations make content easier to understand, and that the relevance of content is demonstrated by real-life examples.* Relevance is a major component of

The POD will provide extensive information for development including many helpful hints and ways to achieve the relevant objective for both in-class and online settings.

As mentioned above, red bars indicate that the faculty member should seek to improve in the relevant objective(s). However, if a faculty member generally earns average to above-average scores and received unusually low scores in a particular course, other course characteristics should be examined to see whether outside variables served to deflate the faculty member’s scores. This information is not only important for interpreting scores for evaluation purposes, but can also help in identifying courses which may need to be upgraded, replaced, etc.

Some characteristics which may be useful to look at would be students’ desire to take the course, student perception of course difficulty, and instructors’ expectations of student responsibility. This information can be found under the Quantitative tab.

Recall the example of a weak summary evaluation reported at the beginning of this document. In that case, the faculty member achieved a score of 3 on his/her summary evaluation. Here is some student and course characteristic data corresponding to that overall score.

Definitely False	More False than True	In Between	More True than False	Definitely True	Mean	Standard Deviation	Did Not Answer	Total Responses
I had a strong desire to take this course.								
8.33% 2	25% 6	20.83% 5	37.5% 9	8.33% 2	3.13	1.13	0	24
I worked harder on this course than on most courses I have taken.								
4.17% 1	12.5% 3	37.5% 9	41.67% 10	4.17% 1	3.29	0.89	0	24
Difficulty of subject matter								
0% 0	0% 0	50% 12	33.33% 8	16.67% 4	3.67	0.75	0	24

One can immediately see that more than half the class expressed “no desire” to “in between” when asked whether they had a strong desire to take this course. This may indicate that it was a required course, perhaps not related to their major. One can also observe that almost half the class felt they worked harder in this course than other courses and that the subject matter was more difficult than other courses, with a large number neutral on the topic as well. That being the case, one would probably not be surprised to see lower scores in a course perceived to be “difficult,” especially when more than half the class did not initially want to take it.

Rather than take an indirect approach to understanding weak scores, we can take a direct approach to see how the students rated the professor and course “on the whole.” We can scroll down to the assessment data for, “Overall, I rate this instructor an excellent teacher,” or “Overall, I rate this course as excellent:”

Definitely False	More False than True	In Between	More True than False	Definitely True	Mean	Standard Deviation	Did Not Answer	Total Resp
Overall, I rate this instructor an excellent teacher.								
12.5% 3	16.67% 4	29.17% 7	37.5% 9	4.17% 1	3.04	1.1	0	24
Overall, I rate this course as excellent.								
8.33% 2	20.83% 5	45.83% 11	16.67% 4	8.33% 2	2.96	1.02	0	24

While the means in both these cases look relatively low, one observes that 17 out of 24 students rated the professor somewhere between average and excellent – even despite the difficulty of the course. Furthermore, while a similar percentage of the class rated the *course* between average and excellent, a much lower percentage felt the *course* was of above average quality than they did the instructor.

Given the above scenario, a chair may discuss the matter with the faculty member, recommend that the faculty member go through relevant [POD IDEA Notes](#) and possibly seek guidance from senior faculty. Furthermore, it would behoove the chair to take a deeper look at the course syllabus and at other course measures in order to gauge course efficacy.

Identifying Course Effectiveness: The “Segment Comparison” Tab

A great way to identify whether a *course* could use some work, is to click on the right-most Segment Comparison tab and see how this section of the course compares with other sections of the same course.

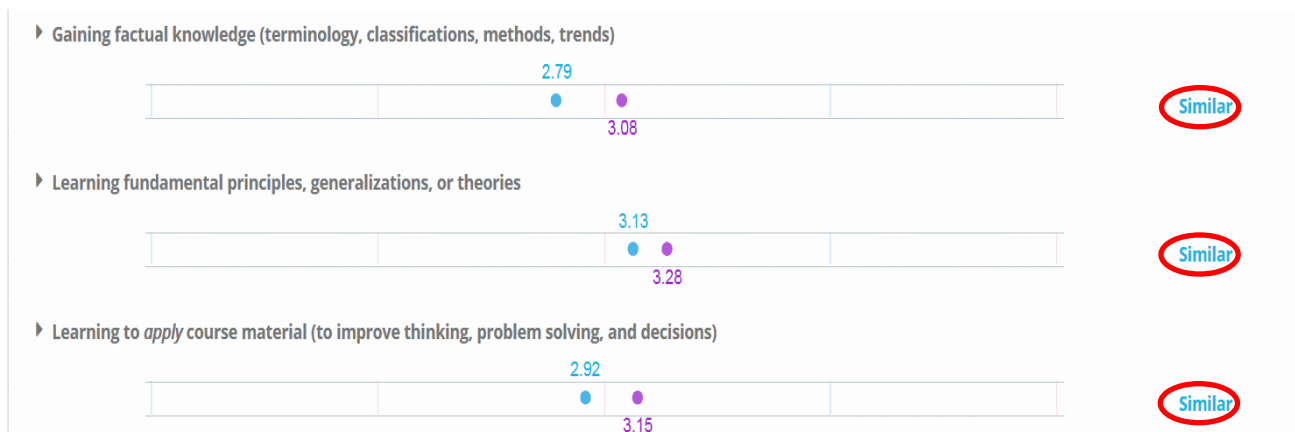


Under the “Segment Comparison” tab, there is a pull-down window so one can look at every question on the survey in comparison to various groups. In this case, we could look at

- 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**

When comparing a particular course to all sections, each objective is once again listed with the faculty member’s scores alongside the scores of other sections of the same course. On the right-hand side in blue is an indicator (**Lower, Similar, Higher**) identifying whether scores of faculty member versus other sections are statistically significantly different.

Here is comparison data corresponding to the original example listed at the beginning of this document, and elaborated upon just above. Recall, the summative evaluation score was equal to 3.



We can see that the faculty members scores are not statistically significantly different from other course sections on all three relevant objectives. This helps to confirm our suspicion that while there may be room for the faculty member to improve, this course probably needs some upgrading. The course setup, structure, and/or curriculum should be analyzed by the department for ways to improve it.

In general, it is important to identify courses that achieve weak scores across the board and try to analyze reasons for it. In some cases, weak scores will simply be due to the fact that a course is very difficult, and perhaps taught to students who are not majors in that field (a highly quantitatively-based course, for example). In other cases, it will be due to the fact that the course is taught in too monotonous a way, or does not offer much new information beyond a course the students have already taken.

Department-Wide Development: A Note

When taking an across-the-board look at departmental courses, chairs and/or departments may notice that certain objectives consistently score low. Some departments may deem it unnecessary to meet certain objectives, for example, the “encouragement of team-work” may not be viewed as a critical objective in statistics classes by that department. In such a case, a low score across-the-board would not be cause for concern. Other objectives, if deemed important, and consistently not satisfactorily met, should be flagged for improvement. The department can discuss where improvements should take place – in a particular course or courses, or across the board in the departmental curriculum. Our companion guide “Using IDEA for Curriculum Assessment and Development,” discusses this topic.

Conclusion

One can see that the “IDEA Evaluation Report,” provides a very rich dataset through which faculty can not only be evaluated but can also be developed and mentored. The data, extensive as it may be, is often very important to analyze in order to properly assess faculty

and course performance, as exhibited in the illustrations above. Summative data does not fully capture the “picture,” nor is it intended to. However, by thoroughly making use of the “Quantitative” and “Segment Comparison” tabs, in tandem with the “Formative” tab and the [POD IDEA Notes](#) offered, areas for both faculty and course improvement can be effectively identified and addressed.