Factsheet for Enhancing Your Use of IDEA Online Course Evaluations

Review carefully this sheet with your feedback report in hand. It gathers key facts you will find useful when planning for and using IDEA (Individual Development and Educational Assessment). To access more information (actual evaluation forms, interpretation guides, related documents) go to the Course Evaluation Program link at http://www.houghton.edu/ira/.

The Diagnostic Form’s Feedback Report (DFR) answers six questions; the Short Form (SFR) omits Q#6:
1) Overall, how effectively was this class taught? (DFR and SFR, page 1)
2) How does this compare with the ratings of other teachers? (DFR and SFR, page 1)
3) Were you more successful in facilitating progress on some class objectives than on others? (FR and SFR, page 2)
4) Do some salient characteristics of this class and its students have implications for instruction? (DFR and SFR, page 2)
5) How did every participant rate every item, including additional questions? (DFR, page 4; SFR, page 3)
6) How can instruction be made more effective? (DFR only, page 3)

Unadjusted (raw) and adjusted scores (scale 1-5): The latter makes classes more comparable by considering four influences beyond the instructor’s control → student’s desire to take this class regardless of instructor, students motivated work habits, class size, and “inherent” course difficulty. Adjusted scores statistically “level the playing field” and make scores comparable for those feeling disadvantaged by teaching unmotivated or poorly prepared students in larger classes.

Average and Converted scores: Average scores [an unfortunate confusing label, in my view, because they are not an “average” of anything] are based on a 5 point scale and are criterion-referenced. Converted scores are standardized (Mean=50; SD=10) and norm-referenced, and account for the fact that some of the 12 objectives (I prefer and use the term outcomes) are easier to achieve and have higher scores. For example, the average student rating of progress for #1 Gaining factual knowledge.... is 4.0 nationwide, while the average rating for classes emphasizing #7 Gaining a broader understanding and appreciation of intellectual/cultural activity .... is 3.69. If only raw averages were considered, instructors choosing certain outcomes could be disadvantaged. Converted scores statistically even out this disparity.

Overall summary index of effective teaching: Rated progress on objectives (A) is the single best estimate of teaching effectiveness. But the full index adds the mean of “excellent teacher” (B) and “excellent course” (C) and weights it the same as A; thus, the overall summary of teaching effectiveness averages A and (B+C/2). See DFR and SFR pg 1 left side.

Example of Criterion-Referenced Index (an institution can determine its own scale, but this is suggested by IDEA):

<table>
<thead>
<tr>
<th>Average Rating</th>
<th>Effectiveness category</th>
<th>Progress Ratings are on a 5 point scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below 3.0</td>
<td>Below acceptable standard</td>
<td>1 = No apparent progress</td>
</tr>
<tr>
<td>3.0-3.4</td>
<td>Marginal, improvement needed</td>
<td>2 = Slight progress</td>
</tr>
<tr>
<td>3.5-3.9</td>
<td>Good</td>
<td>3 = Moderate progress</td>
</tr>
<tr>
<td>4.0-4.4</td>
<td>Excellent</td>
<td>4 = Substantial progress</td>
</tr>
<tr>
<td>4.5 or higher</td>
<td>Outstanding</td>
<td>5 = Exceptional progress</td>
</tr>
</tbody>
</table>

Best practices for optimum and authentic (most accurate and valid) student ratings of instruction:

Prior to semester—
1. On the syllabus place the date (along with all course “due dates”) when students will receive the IDEA email and evaluation link near end of semester (always 14 days prior to the last day of classes). Convey that it is an expectation and is good class citizenship to complete it. Determine whether you will incentivize or just cajole. Our office can provide the list of participants when the window for response is closed.
2. Select 3-5 outcomes as either Essential (double-weighted) or Important (single-weighted) prior to the semester and include these among your syllabus goals or outcomes.
3. IDEA says that when selecting your outcomes, ask these three questions: a) Is this a significant part of the course? b) Do I do something specific to help students accomplish this outcome? c) Does the student’s progress on this outcome affect his or her grade? If yes to all, then include that objective.

During the semester--
4. First day, review syllabus to highlight outcomes/goals and mention that particular assignments are designed to develop specific outcomes (make those connections!). Then reiterate this connection as each assignment comes along.

Near semester’s end--
5. Remind students of your goals/outcomes on the syllabus and how you help them achieve the course outcomes. Keep the big picture in their thinking.
6. Prepare a 3 minute pep talk to motivate their participation. e.g., you really want to know how this course has impacted them, you look forward to their feedback, you use it to make course changes (give an example), you have been assessing them and now it is their turn, mention any incentives, whatever. . . . but not this: “I know it is a pain, but just get it done if you can,” or “Please give me good ratings.” Also, don’t give it the silent treatment and pretend it isn’t happening at all.
(NOTE: Every semester we get emails from students who ask why their form did not come because her roommate got four. We explain our policies about participation, which does not usually satisfy their need to give the professor feedback. Don’t assume that students like it when there is no end of course evaluation.)

Quick hits of crucial things to know:
- The computer ignores (excludes from analysis) all the outcomes marked as Minor. If you check none, IDEA defaults to all 12 as “Important.” On the Faculty Information Form (FIF), mark all twelve as E, I, or M (minor).
- Research by IDEA has demonstrated that selecting 6 or more outcomes has a lowering effect on total ratings. Do not let most or all 12 “ride” for your evaluation. Students apparently do not identify more than about five outcomes as key in your course, in spite of your wishes. Results will very likely be unsatisfactory to you.
- Your converted scores are typically higher when measured against the whole IDEA database compared to disciplinary comparisons or Houghton-only comparisons (DFR and SFR—bottom right of page 1).
- Converted scores have a +/−3.0 margin of error. If you have a measured converted score of 47, your “true” score very likely falls in the range between 44 and 50 (i.e., 47 is not necessarily your actual score).
- Review the full data on the report’s last page to see what other outcomes—the ones you did not select—were scored high by students. Reflect on why: Is a halo effect (all were high—you can do no wrong) going on? Did you spend too much time there and not enough on your intended focus? Should you include that outcome next time?
- If you have, hypothetically, a 3.1 on a selected outcome, review the spread of scores. Could it be for your class of 24 (e.g., their average score ratings 1-5 were, respectively, 6, 5, 1, 3, 9) that this bi-modal score set suggests you may have some ill-prepared students along with very prepared students? Does it have implications for presenting material, assigning group work, or altering course prerequisites?
- To understand the nature of bi-modal ratings of progress, review items 13-15 on the Short Form. Is there evidence of distinct groups who differ in motivation, background, or work habits?
- The minimum goal is always for 75% class participation rate or higher, with 65% being the lowest IDEA suggests using for making assumptions about the true class profile. Ten or more students are ideal, as well. But if it is lower (say, 45%) you can still learn from the empirical responses and especially the free responses.
- IDEA Center says that no more than 30-50% of any personnel decision ought to rely on student evaluations.
- The first two semesters of an instructor’s IDEA use should be treated lightly, due to a learning curve. Note that IDEA never uses an institution’s first year data in their norms.
- IDEA norms are based on nearly 45,000 classes, and thus very stable. Norms for discipline and institution are processed after at least 400 classes are evaluated during the most recent five year period.
- No data from any classes with fewer than 10 respondents are included in the norm groups.
- IDEA’s Technical Report No. 13 shows that there are significant disciplinary differences (particularly STEM compared to others) in “average” ratings. Reasons are unclear: maybe some disciplines draw effective (or ineffective) teachers, or some disciplines are inherently more difficult, complex, or abstract such that they invite especially harsh (or lenient) judgments. Thus there should be no unreflective, direct comparisons across disciplines without regard for these anomalies. Focus should emphasize within discipline comparisons if possible.
- Class size makes a difference: larger classes tend to yield lower ratings than smaller classes. Thus this is part of the adjusted score process.
- Although IDEA ratings have been shown to be reliable (consistent) and valid (useful), all measures of human characteristics have some “measurement noise.” If students rated the class on another day, results might be somewhat different due to sampling errors, but not much. For classes in the 15-34 student range, a sampling error of +/−0.2 (on a 5 point scale) is typical. It is slightly higher for smaller classes and lower for larger classes.
- If you need electronic forms of IDEA reports for the Rank and Tenure Committee process, ask our office for help.
- Do students more readily respond to the Short Form than the longer Diagnostic Form? No and here is the proof:

<table>
<thead>
<tr>
<th>Survey Type</th>
<th>Number of Classes</th>
<th>Number Enrolled</th>
<th>Number Responded</th>
<th>Average Class Response Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diagnostic</td>
<td>636</td>
<td>11,112</td>
<td>8,428</td>
<td>76.6%</td>
</tr>
<tr>
<td>Short</td>
<td>383</td>
<td>6,830</td>
<td>4,760</td>
<td>71.1%</td>
</tr>
<tr>
<td>All Classes</td>
<td>1,019</td>
<td>17,942</td>
<td>13,188</td>
<td>74.6%</td>
</tr>
</tbody>
</table>

Office of Institutional Research and Assessment -- August, 2012