

In our continuing efforts to provide institutions with the best information possible, changes were made in 2004 in the way we calculate the NSSE benchmarks of effective educational practice. These changes allowed us to produce student-level benchmark scores, enhancing the usability of the information for intra-institutional comparisons. For example, institutions can now examine benchmarks at the school, college, or department level, or can compare particular subgroups of students (e.g., men and women or seniors from two different years). The changes in the calculation require that benchmarks prior to 2004 be recalculated to more accurately compare institutional performance over the years using the same metric (Table 1).

Another change made to the survey in 2004 affects the information in Table 1. Response options for the ‘enriching’ items (question 7 on the survey) were altered in 2004 making it untenable to compare newer results on these items with those of 2003 and earlier. For this reason, the Student-Faculty Interaction benchmark is recalculated without one item and the Enriching Educational Experiences benchmark is not recalculated.

**Table 1**  
**Recalculated Benchmarks for NSSE Participation since 2001 <sup>a</sup>**

<i>Benchmark</i>	<i>Class</i>	<i>2001</i>	<i>2002</i>	<i>2003</i>	<i>2004<sup>b</sup></i>	<i>2005</i>	<i>2006</i>	<i>2007</i>
Level of Academic Challenge	FY		49.1			49.5		50.6
	SR		54.9			54.0		54.8
Active and Collaborative Learning	FY		40.2			40.8		42.2
	SR		49.0			49.4		51.5
Student-Faculty Interaction <sup>c</sup>	FY		35.8			36.2		36.8
	SR		43.2			43.1		45.4
Supportive Campus Environment	FY		56.3			54.6		54.2
	SR		49.7			50.0		52.0

For more information about benchmark construction and to download syntax that calculates student-level scores, visit the NSSE 2007 Institutional Report Web site: [www.nsse.iub.edu/2007\\_Institutional\\_Report](http://www.nsse.iub.edu/2007_Institutional_Report).

## How comparable are benchmark scores from year-to-year?

This report is a brief introduction to comparing institutional performance over time, not an exhaustive treatment of all the pertinent issues that need to be considered. We recommend that you do further analysis to better understand the changes within your institutional context. It is important to keep in mind three issues before comparing benchmark scores from year-to-year:

- 1) Drawing a random sample from a population results in a certain amount of sampling error – an estimate of the degree to which the characteristics of the sample do not match those of the population. Smaller samples relative to the size of the population risk larger sampling errors. Thus, relatively small benchmark differences could be attributed to random sampling fluctuation.
- 2) In addition to sampling error, you should examine the demographic characteristics of the samples to be sure that similar groups of students are represented among the respondents in various years. If respondent characteristics are different, and these differences could likely affect engagement scores, they should be acknowledged and taken into account when attributing reasons for benchmark differences. A more sophisticated approach would be to weight the samples so they more closely resemble the student population, and then recalculate the benchmark scores using the formulas provided by NSSE. However, keep in mind that all of your recalculated benchmarks are weighted by gender and enrollment status.<sup>b</sup>
- 3) Some questions and response options were changed over the years based on psychometric analyses to

improve the survey's validity and reliability. Most notably, response options for the 'enriching' items (question 7 on the survey) were revised in 2004.<sup>d</sup> Our analysis shows that these items are not comparable with prior years. *For most institutions, this change will produce a substantially lower Enriching Educational Experiences score since 2004 compared to prior years, particularly for first-year students.*

## What constitutes a real change in a benchmark score?

One way to estimate the magnitude of change in a benchmark score over time is to combine your institutional data from all participating years and run statistical analyses between students from the respective years. For example, t-tests can be computed between first-year students in 2003 and first-year students in 2006 to see if the differences between benchmark scores are statistically significant. Effect sizes can also be computed by dividing the difference of the benchmark scores by the standard deviation of the entire distribution. The t-tests can also be weighted according to statistical weights provided by NSSE (based on gender and enrollment status), or institutions can create their own weights based on school records.

Institutions can also conduct regression analyses using the multi-year data and include a dummy variable for the year of participation as an independent variable. With this approach, the regression model could control for student demographic variables or other independent variables to see what the unique effect of the year of administration might be.

## Notes

- |  |  |  |
|--|--|--|
| <p>a. Scores from NSSE 2000 are not included because several significant changes were made to the survey instrument after that year, thus making year-to-year comparisons less suitable.</p> <p>b. Student weights prior to 2004 were computed exclusively using the most recent IPEDS data available. Starting with 2004, institutional population files were used for class rank and gender because these files provide more recent and accurate data. Beginning in 2005, enrollment status information (full-time/part-time) was also</p> | <p>taken from institutional population files rather than IPEDS.</p> <p>c. All items in question 7 on the current NSSE instrument were rescaled in 2004. One of these items, "Work on a research project with a faculty member outside of course or program requirements," contributes to the Student-Faculty Interaction benchmark. See note 'd' for more details. Therefore the Student-Faculty Interaction scores on this report do not include the 'research' item. This also means that the score on this report will not match benchmarks reported on</p> | <p>previous year reports, or on your 2007 Benchmark Comparisons report.</p> <p>d. All items in question 7 on the 2004 instrument were rescaled in 2004. The old response set (NSSE 2000-2003) was 'yes,' 'no,' or 'undecided' whereas the new (NSSE 2004-2007) response set is 'done,' 'plan to do,' 'do not plan to do,' or 'have not decided.' Our analysis shows that these items are not comparable across years. Therefore, it is not possible to compare the 2004-2007 Enriching Educational Experiences benchmark with prior years (2001-2003).</p> |
|--|--|--|