

In 2004, changes were made in the process for calculating NSSE Benchmarks of Effective Educational Practice scores. The changes are a result of our continuing efforts to provide institutions with the best information possible. By revising our calculation process, we have enhanced the usability of the information for within institutional comparisons. For example, institutions can calculate benchmark-like scores at the school, college, or department level. This was not previously possible, because the benchmarks were constructed using institution-level data. In addition, using the student-level scale scores, the pre-cursors to the benchmarks, institutions can compare groups of students (e.g., seniors from two different years). For more information about the benchmark construction process, examples of possible analyses, and descriptions of how to calculate student-level scale scores, please see the NSSE 2004 annual report website.

While individual institutions now have more options to reconstruct NSSE benchmark scores for their own purposes, the changes in the benchmark calculation procedures require that benchmarks from previous years also be recalculated in order to accurately interpret changes in institutional performance over the years. This report provides NSSE 2004 and prior year benchmark scores based upon the revised process. Thus, the information below allows you to compare benchmark scores from two or more years using the same metric.

Recalculated Benchmarks

The table below shows first-year and senior recalculated benchmarks for each year of NSSE participation since 2001.^a The benchmarks for previous years are produced using the improved calculation method developed for use in 2004. Note also that the 2004 Student Faculty Interaction benchmark^d has been computed in a way to make possible accurate year-to-year comparisons.

Recalculated NSSE Benchmarks^a		2001 ^b	2002	2003	2004 ^c
Level of Academic Challenge	FY		56		55
	SR		61		63
Active and Collaborative Learning	FY		47		46
	SR		54		56
Student-Faculty Interaction ^d	FY		48		42
	SR		54		58
Enriching Educational Experiences ^e	FY		56		30
	SR		51		50
Supportive Campus Environment	FY		70		65
	SR		70		66

How comparable are benchmark scores from year to year?

First, this report is a brief introduction to how to compare 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 and investigation to better understand the changes in relation to your institutional context. It is important to keep in mind the following 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 likely could affect engagement scores, these differences 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.
- 3) Some questions and response options were changed over the years after 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. Our analysis shows that these items are not comparable with prior years. *For most institutions, this change will produce a substantially lower score in 2004 compared to prior years, particularly for first-year students.* See the NSSE website for specific changes to these and other items.

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 2004 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 this 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.

For detailed examples of these procedures, please see the 2004 Annual Report section of the NSSE website (www.iub.edu/~nsse).

Notes

- ^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. Details about survey changes from year to year can be found on the NSSE website: www.iub.edu/~nsse.
- ^b In 2001, only the norm sampled students were included in institutional benchmark scores. In 2002 – 2004, all randomly sampled students are included.
- ^c Student weights prior to 2004 were computed exclusively using the most recent IPEDS data available. In 2004, institutional population files were used for class rank and gender because these files provide more recent and

- accurate data. IPEDS data were still used in 2004 for enrollment status information (full-time/part-time).
- ^d All items in question 7 on the 2004 instrument were rescaled on the 2004 survey. 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. The old response (NSSE 2000-2003) set was "yes, no, or undecided" whereas the new response (NSSE 2004) 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 the Student-Faculty Interaction scores on this report do not include

- the "research" item. All items comprising the benchmark on this report are identical, so that 2004 results can be compared with prior years. This also means that the score on this report will not match benchmarks reported on previous year reports.
- ^e All enriching educational experience items (question 7 on the 2004 instrument) were rescaled on the 2004 survey (see 3 above). Also, the "learning community" item (question 7c on the 2004 instrument) was not on the 2001 instrument and is therefore not part of the 2001 EEE benchmark.