NSSE 2008
Psychometric Properties
The National Survey of Student Engagement (NSSE) was designed to assess the extent to which students participate in empirically-derived effective educational practices and what they gain from their college experience. A large, growing body of research on college student development shows that the time and energy students devote to educationally purposeful activities contributes to their learning and personal development (See NSSE Conceptual Framework at www.nsse.iub.edu/html/researchers.cfm for more details). NSSE collects data on student behaviors that are highly correlated with many desirable learning and personal development outcomes of a college education. This document summarizes many of the projects that the NSSE research team conducts in order to measure various psychometric properties of NSSE, beginning with an overview of the content and construction of the survey instrument. It also discusses various measurements of validity and reliability as well as investigations of potential bias. This document concludes with information on where to find additional psychometric NSSE information.

Validity
The validity of a survey refers to how well the survey measures what it is intended to measure. This section summarizes many of the ways the NSSE research team analyzed the survey instrument’s validity: through question creation, question analysis, and correlations with various student outcomes.

What does the instrument cover?
NSSE asks students to report how often they participate in activities that represent good educational practice. The survey also covers students’ perceptions of the college environment associated with achievement and satisfaction. In addition, students are asked to estimate their educational and personal growth since starting college. Finally, students provide information about their background, including age, gender, race or ethnicity, living situation, educational status, and major.

Does the instrument yield valid information?
The NSSE research team worked diligently to ensure that survey items were clearly worded, well-defined, and had high content and construct validity. Cognitive interviews and focus groups revealed that students had difficulty interpreting as intended very few of the survey items. Although some students had trouble understanding such things as the meaning of a learning community and had difficulty determining the difference between socializing and relaxing, these problems were consistent across different types of students from different types of institutions. Additionally, items that contribute to the five NSSE benchmarks were not problematic, implying that the benchmarks are also valid measures of the quality of student engagement experiences.

In the Connecting the Dots report (www.nsse.iub.edu/pdf/Connecting_the_Dots_Report.pdf) researchers used qualitative methods to investigate whether or not NSSE survey questions were working as intended for different types of students at different types of institutions. The researchers found that the NSSE survey works equally well for students from different racial and ethnic backgrounds as well as for students from different types of institutions.

Overall, the pattern of responses from first-year students and seniors suggests the items measure what they are supposed to measure. For example, as one would expect, seniors are, on average, more engaged in educational pursuits that involve working on research with faculty members, tutoring other students, and talking about career plans with an advisor. Senior students are likely to be further in their program of study and more likely to be planning for their futures after graduation. First-year students are, on average, more engaged in educational pursuits such as preparing two or more drafts of a paper, participating in co-curricular activities, and taking part in experiences that help them to understand people of other racial and ethnic backgrounds.

First-year students are more likely to take classes that require multiple drafts of papers, or seniors may need fewer drafts of papers to present acceptable work. First-year students are also more likely to live on campus which puts them in closer proximity to co-curricular
activities and peers from different backgrounds. These
differences in responses to items are not surprising and
support the validity of the NSSE survey instrument.

How does student engagement relate
to other student outcomes?
The NSSE survey includes a number of self-reported
student outcome measures such as educational and
personal growth, average grades, and satisfaction. An
exploratory factor analysis based on all randomly sampled
students who responded to the NSSE 2006 educational
and personal growth items in question 11 yielded three factors:
personal and social development, practical competence,
and general education. NSSE also uses a satisfaction scale
comprised of answers to question 12 that asked students to
evaluate their entire educational experience and
question 13 that asks whether students would attend the
same institution again if they could start over. The table
above shows the correlations between NSSE benchmarks
of effective educational practice and these self-reported
outcomes based upon NSSE 2007 data.

More details about student engagement and college
outcomes can be found in the Connecting the Dots report.
In this report, researchers found that student engagement
during college had a positive effect on students’
first-year grades and persistence to the second year of
college while controlling for a variety of pre-college and
first-year experience variables such as pre-college GPA
and number of hours per week working off-campus.
Although student engagement during college is a benefit
for students of all racial and ethnic backgrounds, this study
found that for historically underserved students, the gains
may be greater. For example, increases in student
involvement resulted in higher gains in GPA for Hispanic
students than White students. Similarly, African American
students and female students engaging in educationally
meaningful activities were more likely to persist to their
second year of college than White students and male
students, respectively.

Can we trust student
self-reported data?
The credibility of self-reports have been examined
extensively. Self-reported data is likely to be valid under
five general conditions: (1) when the information
requested is known to the respondents; (2) the questions
are phrased clearly and unambiguously; (3) the questions
refer to recent activities; (4) the respondents think the
questions merit a serious and thoughtful response; and (5)
answering the questions does not threaten, embarrass, or
violate the privacy of respondents or encourage
respondents to respond in socially desirable ways
(Bradburn & Sudman, 1988; Brandt, 1958; Converse &
Presser, 1989; DeNisi & Shaw, 1977; Hansford & Hattie,
1982; Laing, Swayer, & Noble, 1989; Lowman &
Williams, 1987; Pace, 1985; Pike, 1995). NSSE was
intentionally designed to satisfy all these conditions.

How often is often?
Survey researchers often wonder about the meaning of
vague quantifiers such as “sometimes” or “often” as
employed by the NSSE survey. When we use results
from these questions in our assessment efforts and
research, we assume that the following questions can all
be answered affirmatively:

- Does each response option have a distinct
  meaning (e.g., Does “often” mean something
different than “sometimes”)?
- Do the assumed intervals between the
  options progressively increase in frequency
  from “never” to “very often”?
- Are the intervals approximately equal (e.g.,
  “very often” means nine times per week, “often”

<table>
<thead>
<tr>
<th>NSSE Benchmarks</th>
<th>Practical Competence</th>
<th>General Education</th>
<th>Personal Social</th>
<th>Grades</th>
<th>Satisfaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of Academic Challenge</td>
<td>.48 .44</td>
<td>.49 .47</td>
<td>.42 .39</td>
<td>.16 .13</td>
<td>.26 .25</td>
</tr>
<tr>
<td>Active &amp; Collaborative Learning</td>
<td>.40 .39</td>
<td>.35 .33</td>
<td>.38 .35</td>
<td>.14 .15</td>
<td>.23 .21</td>
</tr>
<tr>
<td>Student Faculty Interaction</td>
<td>.39 .36</td>
<td>.35 .34</td>
<td>.41 .38</td>
<td>.07 .16</td>
<td>.22 .26</td>
</tr>
<tr>
<td>Enriching Educational Experiences</td>
<td>.34 .28</td>
<td>.29 .27</td>
<td>.36 .34</td>
<td>.10 .16</td>
<td>.20 .20</td>
</tr>
<tr>
<td>Supportive Campus Environment</td>
<td>.57 .56</td>
<td>.53 .52</td>
<td>.57 .57</td>
<td>.09 .12</td>
<td>.55 .58</td>
</tr>
</tbody>
</table>

Note: All correlations are significant at the p < .01 level.
means six times per week, and “sometimes” means three times per week)?

- Do response options change their meaning from item to item (e.g., “often” asking questions in class meant doing so six times per week, whereas “often” discussing ideas outside of class meant doing so twice per week)?

In 2006, we asked students to quantify their responses to several survey items to which they responded with vague quantifiers earlier on the survey. The results show that across the board students on average assigned distinct and increasing quantities to “never,” “sometimes,” “often,” and “very often.” For example, when asked how often they asked questions in class or contributed to class discussions, students said that “never” meant 0 to 1 times per week, “sometimes” meant 2 times per week, “often” meant 6 times per week, and “very often” meant 15 times per week. As this example shows, we found that for most items the intervals between response options are roughly even (see figures at right). Additionally, we found that students adapted the meaning of the vague response options from item to item. In the figures, for example, “very often” means 15 times per week for one item and only 5 times per week for the other.

Reliability

Student responses to the survey are reliable to the extent that they are consistent and reproducible. Research analysts at NSSE examined the reliability of student responses in two ways: test-retest analysis at the student level and stability analysis at the institution level.

How stable are students’ responses between survey administrations?

Assuming little variation in an individual student’s behavior within a short time period, we expect consistent or reliable responses to the survey items. In 2002, we conducted a test-retest analysis using 1,226 respondents who completed the same form of the paper survey twice over a period of several months. For the students’ responses on the items related to three of the benchmarks (i.e., level of academic challenge, active and collaborative learning, and enriching educational experiences), the reliability coefficients were 0.74. Student responses for the items related to student-faculty interaction and to supportive campus environment had reliability coefficients of 0.75 and 0.78, respectively. In 2005, we conducted the study again using 1,536 respondents who completed the paper or Web survey twice within a period of several months. The results were similar to the earlier study with the reliability coefficients ranging from 0.69 (level of academic challenge) to 0.74 (enriching...
educational experiences). The following table shows the test-retest analysis results from the 2002 and 2005 NSSE survey administration. These findings suggest little variation in student responses from one testing period to the next.

<table>
<thead>
<tr>
<th>NSSE Benchmarks</th>
<th>2002</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of Academic Challenge</td>
<td>0.74</td>
<td>0.69</td>
</tr>
<tr>
<td>Active and Collaborative Learning</td>
<td>0.74</td>
<td>0.72</td>
</tr>
<tr>
<td>Student-Faculty Interaction</td>
<td>0.75</td>
<td>0.70</td>
</tr>
<tr>
<td>Enriching Educational Experiences</td>
<td>0.74</td>
<td>0.74</td>
</tr>
<tr>
<td>Supportive Campus Environment</td>
<td>0.78</td>
<td>0.70</td>
</tr>
<tr>
<td>N</td>
<td>1226</td>
<td>1536</td>
</tr>
</tbody>
</table>

How stable are institutions’ scores between survey administrations?

Assuming no major shifts in an institution’s policies, we would expect an institution to have relatively stable or reliable benchmark scores from one year to the next. In 2003, we conducted a stability analysis to measure the strength of the associations between benchmark scores for 214 institutions that participated in the 2002 and 2003 administrations of the survey. The benchmark scores were calculated using unweighted student responses to survey items that were similar for the two years. Values of the Spearman’s rho correlations for these benchmark scores ranged from 0.81 (student-faculty interaction) to 0.88 (level of academic challenge) for first-year students, and from 0.83 (active and collaborative learning) to 0.93 (enriching educational experiences) for seniors. We conducted this study again using data from 236 institutions that participated in both the 2004 and 2005 NSSE survey administrations. The results of the study showed the Spearman’s rho correlations ranged from 0.78 (student-faculty interaction) to 0.89 (enriching educational experiences) for first-year students, and from 0.78 (active and collaborative learning) to 0.92 (enriching educational experiences) for seniors. These findings suggest that institution-level NSSE data are relatively stable from year to year.

Do non-respondents differ from respondents?

Psychometric bias refers to a poor estimate of true scores in a population due to variants such as respondent characteristics or testing situations. The NSSE research team has investigated potential bias in a variety of ways including analysis of non-response, mode of administration, type of institution, and students’ ethnicity. To determine whether respondents and non-respondents differed in their engagement in selected effective educational practices, the Indiana University Center for Survey Research conducted telephone interviews with 553 non-respondents from 21 different colleges and universities that participated in the NSSE 2001 survey administration. A similar study was conducted again in the NSSE 2005 survey administration with 1,400 non-respondents from 24 different colleges and universities. In addition, we also conducted a nonresponse study by comparing NSSE 2005 benchmark scores of early and late respondents. Although some differences were found between respondents and non-respondents, no consistent trend was found to support the existence of non-response bias. Generally speaking, undergraduate students who do not complete the NSSE survey when invited to do so may actually be slightly more engaged than respondents. This is counter to what many observers believe, that non-respondents have a less educationally productive experience and, as a result, do not respond to surveys. The results of the non-response and early-late respondent study show no significant sign of non-response bias in NSSE.

Do students respond differently depending on the mode of administration (paper vs. Web)?

Using ordinary least squares (OLS) we analyzed NSSE 2000 data to ascertain whether students who completed the survey on the Web responded differently than those who responded via a traditional paper format. We controlled a variety of student and institutional characteristics that may be associated with either engagement or mode. Responses to Web and paper surveys showed small, but consistent, differences that tended to favor the Web mode where differences existed. Items related to computing and information technology exhibited some of the largest effects favoring the Web, which is not surprising, given that many students who receive a paper survey chose to complete it using the Web option, suggesting a predilection for technology. On the other hand, students who answered paper surveys spent more time preparing for class and did more reading and writing. These findings, combined with previous analysis, especially for items unrelated to computing and information technology, are generally consistent with the results from single institution studies. The full-length report can be downloaded from: [www.nsse.iub.edu/pdf/mode.pdf](http://www.nsse.iub.edu/pdf/mode.pdf)

The percentage of students who respond to NSSE using the Web version has increased dramatically over the years. In 2000, less than 40% of the respondents completed the Web version. By 2007, more than 95% of the respondents completed the survey via the Web.
Because the vast majority of NSSE respondents use the Web version, the two forms of the survey have even less of an effect on an institution’s results.

**Where can we find additional psychometric information on NSSE?**

NSSE has a growing portfolio of psychometric tests and analyses that it conducts on a regular basis. A more comprehensive summary of this work can be found at the NSSE Web site at [www.nsse.iub.edu/html/researchers.cfm](http://www.nsse.iub.edu/html/researchers.cfm).

**References**


