EXAMINING THE THREAT OF NONRESPONSE BIAS AS A FOLLOW-UP TO THE NATIONAL SURVEY OF STUDENT ENGAGEMENT

Debra Allen
Assistant Director
Office of Institutional Research
University of Maine

Theodore Coladarci
Director
Office of Institutional Research
University of Maine

The National Survey of Student Engagement (NSSE), an online survey administered to freshmen and seniors, can be used to measure the quality of student experiences. NSSE survey items, described by NSSE as representing "empirically confirmed good practices," relate to individual student behaviors as well as student perceptions of their college experience. Institutions—over 750 last year—use NSSE to support decision making, design goals, and analyze progress in such areas as accreditation, accountability reporting, strategic planning, and program assessment (Banga, Pike, & Hansen, 2009).

NSSE can be a valuable tool, but with a national response rate of approximately 27% and roughly one third of institutions achieving response rates below 30% (NSSE, 2011), there are concerns regarding the degree to which institutions can generalize their results to their student population. Low response rates do not necessarily suggest a lack of representativeness, however; a survey yielding a low response rate may still produce a sample that is representative of its population (Dey, 1997; Groves, 2006). It is only when there are discrepancies between responders and nonresponders in perceptions or behaviors relevant to the survey topic that nonresponse bias is introduced and becomes a threat to the validity of inferences made from the survey results. NSSE acknowledges the impact that nonresponse bias may have on the interpretation of an institution's results and therefore encourages institutions to conduct their own nonresponse studies (Chen, 2006).

In short, our purpose in conducting the present study was to meet NSSE's call for institutional-level nonresponse studies. Our institution, the University of Maine, is a land-grant university with a total student population of approximately 11,000 (roughly 8,270 of whom are degree-seeking undergraduates). UMaine administered the web-based NSSE to all freshmen and seniors in spring 2011, achieving a response rate of approximately 21%. To determine the

degree to which these respondents are representative of the remaining 79% of freshmen and seniors, we addressed the following questions:

a) How do the demographic characteristics of nonresponders compare with those of the responders?
b) What reasons for nonresponse are subsequently offered by nonresponders?
c) How do nonresponders' perceptions of their UMaine experience compare with those of responders, using a selection of items from the NSSE survey?

Related Literature

NSSE Overview

NSSE emerged from the efforts of a design team, lead by Peter Ewell of the National Center for Higher Education Management Systems, that had been tasked by Pew Charitable Trusts to develop an instrument to measure the extent to which college students show good educational practices and to assess what they gain from their college experience (Kuh, 2009). The survey was first administered to 276 institutions in 2000; in 2011, 751 institutions from the United States and Canada participated (NSSE, 2011). Institutions use NSSE to identify areas where institutional policies or practices may be improved to promote good educational practices, and to serve as an external accountability measure of overall quality (NSSE, 2011). The Voluntary System of Accountability (VSA), an initiative through which institutions report measures of student outcomes, chose NSSE as one of the options schools can use for assessing student engagement. NSSE data, along with other information provided in the spirit of accountability and public disclosure, appear conspicuously on the participating institution's website, in the "College Portrait." In addition to engagement data, the College Portrait provides general consumer information as well as measures of student success and progress, learning outcomes, and perceptions and experiences.

The three core purposes of the NSSE are to provide actionable data to institutions so that they can improve the undergraduate experience, highlight and document effective educational higher education practices, and advocate for public acceptance and use of standard measures for college quality (Kuh, 2009). Through five sets of questions, students are asked about (a) their participation in sports and activities, study time, interaction with faculty members; (b) what they see their institution as requiring of them; (c) their perceptions of their college environment (including factors believed to impact satisfaction, academic achievement, and persistence); (d) information regarding their socioeconomic and educational background; and e) their educational and personal growth since starting at the institution.

For ease in analysis and interpretation, NSSE developed five benchmark scales of effective educational practice that capture institutional characteristics and student behavior highly aligned with both learning and personal development. Comprising a total of 42 items, the five benchmark scales are Level of Academic Challenge, Active and Collaborative Learning, Student-Faculty Interaction, Enriching Educational Experiences, and Supportive Campus Environment (Kuh, 2009).

Although NSSE is a well-known and frequently used survey, validity concerns have been raised regarding the accuracy of student self-reports (Porter, Rumann, & Pontius, 2011), the weak relationship with student academic outcomes (Gordon, Ludlum, & Hoey, 2008), and low response rates (Porter & Whitcomb, 2005). In light of these concerns, institutions should take steps to verify the validity of the data collected at their own institution. In fact, Chen et. al. (2009), in their outline of best practices for analyzing NSSE data, suggest that determining the quality of an individual institution's data should be a priority for any analyst using NSSE data. They stress the importance of verifying that population estimates are accurate, and recommend

that institutions consider sampling error, analyze the potential for nonresponse bias, and examine the proportional representation of student subgroups within their responder samples.

Nonresponse bias in survey research

Over time, survey response rates have suffered a noticeable decline (Dey, 1997). As a result, the awareness of and concerns regarding nonresponse bias has also grown in recent years. The following is a more comprehensive definition of nonresponse bias and a summary of the methods used to examine its existence.

Definition of nonresponse bias. As mentioned above, the literature on nonresponse bias suggests that bias is not necessarily a function of low response rates. Rather, it is only when there is a discrepancy between respondents and nonrespondents on relevant attitudes, perceptions, and behaviors that bias is introduced. Low response rates then serve to exacerbate the level of bias (Groves, 2006; Pike 2008).

Nonresponse bias, and the degree to which it is a problem, is often related to the reasons for nonresponse (Groves, 2006; Rogelberg & Luong, 1998). Rogelberg and Luong (1998) outline the four major classes of nonresponse as inaccessibility, inability to respond, carelessness, and noncompliance. Each reason is likely to affect bias in a different way, something that Groves (2006) stresses should be taken into account by survey researchers. For example, nonresponse due to an inability to contact a person is likely to introduce a different level of bias than nonresponse because of a refusal to answer questions about the particular subject. Both reasons will have an impact on response rate, but biased responses are more likely with the latter. The former may be completely unrelated to the attitudes, behaviors, or perceptions of interest in the survey.

Thus, there are four important considerations when interpreting survey results: response rates across various subgroups in the population of interest; demographic differences between

responders and nonresponders; distinctions between responders and nonresponders in the attitudes, perceptions, or behaviors measured by the survey; and reasons for nonresponse. The first two can be examined using existing data that are often readily available, but one must collect data from the nonresponder group in order to address the latter two concerns.

Methods for examining nonresponse bias. Nonresponse bias can be examined using a number of methods. There are three approaches researchers regard most promising for studying nonresponse (Dey, 1997; Groves, 2006; Hartman, Fuqua, & Jenkins, 1986; Porter & Whitcombe, 2005; Rogelberg & Luong, 1998). In the *archival approach*, one compares nonresponders and responders on a number of relevant demographic variables and then examining each variable's relationship with the survey responses. In a *wave analysis*, one compares early and late responders, assuming that late responders are most similar to nonresponders. And in a *follow-up analysis*, one obtains responses from a sample of nonresponders and then examines the degree to which the two groups differ on survey items of interest.

These three approaches have both strengths and weaknesses. For example, although the analysis of demographic variables used in the archival approach can be conducted with a simple matching of data, the information provided only gives a partial indicator of the extent to which bias exists; because it does not provide any information on nonresponders (Groves, 2006). A wave analysis is also relatively easy to conduct, particularly if the survey was administered through a process of multiple reminders, but it assumes that nonresponse bias will be systematically distributed over time (Hartman, Fuqua, & Jenkins, 1986). Because of these considerations, the follow-up analysis is a common approach to understanding nonresponse. Its primary benefits are that it allows for analyses of the attitudes and behaviors of both responders and nonresponders (Porter & Whitcombe, 2005) and therefore can provide a basis for estimating their differences—and, therefore, the extent of any bias (Hartman, Fuqua, & Jenkins, 1986).

That said, the follow-up approach is not without its weaknesses. Its primary weakness (ironically) is the risk of additional bias introduced with the choice of method for data collection. If the method used to collect responses from the follow-up sample distorts who responds and the answers they furnish, the responses will not provide an unbiased view of nonresponders (Dey, 1997; Porter & Whitcombe, 2005; Rogelberg & Luong, 1998). For example, telephone surveys may produce biased results (Dillman, Sangster, Tarnai, & Rockwood, 1996). Having reviewed the literature on the comparisons of responses to mailed or phone surveys, Dillman et. al. propose that telephone interviews are more likely than mail questionnaires to produce (a) socially desirable and acquiescent answers, (b) question-order effects, (c) quick answers that reflect a general standard held by the respondent, and (d) extremeness on response scales. These limitations call for caution when interpreting follow-up studies conducted through telephone surveys.

Analyses of nonresponse bias in NSSE Surveys

NSSE researchers have used both the archival and follow-up approaches to examine nonresponse.

Archival approach. NSSE researchers have consistently found significantly lower response rates among men and part-time students than women and full-time students (NSSE, 2011). NSSE uses a weighting procedure in their institutional report to compensate for such differences.

In their multi-level study, Porter and Umbach (2006) used the NSSE to examine how response rates vary by the makeup of the student body and institutional characteristics. With a sample of approximately 167,000 students across 321 schools, these researchers also found women were more likely to respond than men. In addition to differences associated with gender,

their results showed that student ability and such social environment factors as density, urbanity, and the percentage of part-time students also affected institutional response rates.

NSSE (2008) examined the relationship between levels of high school engagement, as measured by the Beginning College of Student Engagement, and whether or not a student responded to the NSSE in the spring of the freshman year. Based on data from approximately 35,000 students across 89 institutions, the analyses revealed no relationship between high school engagement and propensity to respond.

Follow-up studies. Two national NSSE follow-up studies suggest the existence of nonresponse bias. In a 2001 study, the Indiana Center for Survey Research (CSR) conducted a nonresponse analysis based on follow-up telephone surveys with 553 nonresponders from 21 institutions (Kuh, 2003). The interviews included 21 engagement and 3 demographic items from the survey. Freshmen nonresponders scored higher than respondents on nine items, while responders only scored higher on three items. Senior nonresponders scored higher than respondents on six items, and responders scored higher on the same three items seen in the first-year group. In general the results showed a slightly higher level of engagement among nonresponders than responders (although the CSR researchers acknowledged the need for caution in interpreting their results due to the potential bias introduced by the use of a telephone interview).

We see somewhat similar results in a later study, also conducted by the CSR (NSSE, 2006). The second study included phone interviews with 1,408 nonresponders from 24 different institutions. The telephone interviews included 17 questions, with items representing student-faculty interaction, the campus environment, and developmental-gain subscales. Four demographic items also were included. The results of the telephone interviews suggest nonresponders were more likely than respondents to view faculty, staff, and their campus as

supportive. However, these two groups did not differ in student-faculty interaction or developmental gains.

McInnis (2006) used an additional mailing of the survey itself rather than a telephone survey to reach first-year nonresponders. McInnis received surveys from 25 of 94 nonresponders (26.6% response rate). Similar to the previous two NSSE studies conducted by CSR there were minimal differences. The only scale that showed a significant difference was the faculty interaction scale, with nonrespondents showing higher mean scores than respondents.

The results of the national nonresponse bias studies conducted by CSR and the small-scale study conducted by McInnis (2006) suggest there is potential for nonresponse bias to threaten the generalizability of NSSE results. Although each study suggests a slightly higher level of engagement in some areas among nonresponders, the threat of bias associated with NSSE's use of follow-up telephone surveys and the small scale of the McInnis study do not provide strong evidence that the same differences may be present at all schools. Again, the degree to which such a bias exists is likely to differ across institutions. As we consider the results of the 2011 NSSE results for UMaine, it therefore is helpful to know how responders may differ from nonresponders, in terms of both their demographic characteristics and their perceptions of the UMaine experience.

Method

As reported early the three research questions we addressed in this study of nonresponse bias are:

(a) How do the demographic characteristics of nonresponders compare with those of the responders?

(b) What reasons for nonresponse are subsequently offered by nonresponders?

(c) How do nonresponders' perceptions of their UMaine experience compare with those of responders, using a selection of items from the NSSE survey?

We used the archival approach to answer the first question and a follow-up analysis to address the second and third questions.

Archival approach. We began by matching NSSE's list of responders and nonresponders to UMaine's spring 2011 student database. We then created a dataset comprising these student characteristics: gender, residency (Maine resident, nonresident), enrollment status (part-time, full-time), cumulative GPA (below 2.5, 2.5-2.9, 3.0-3.4, 3.5 and above), transfer status (new student, transfer student), living situation (on-campus, off-campus), and college of major. We used the χ^2 test of independence to determine whether there was a relationship between NSSE participation and the respective demographic variables.

Follow-up analysis. To assess the perceptions of NSSE nonresponders, we conducted brief telephone interviews with 50 freshmen and 50 seniors whom we randomly sampled from the list of UMaine's nonresponders. Its possible biases notwithstanding, we chose this method because of the short timeframe it permitted and its relatively inexpensive cost.

To contact the students, we pulled random samples of approximately 50 freshman and 50 senior names and telephone numbers from the nonresponder list provided by NSSE. We called students during both day and evening hours to reach working and non-working students alike. In the event that a student was not home, we did not leave a message (to relieve the student of the burden of calling back). We attempted to contact each student in our initial sample three times. Once we had either reached or exhausted three attempts to contact each student, we repeated the process with three additional random samples of 50 freshmen and 50 seniors. In total, we attempted to reach approximately 370 students. The data collection occurred between June 1, 2011 and August 11, 2011.

To keep the interviews to two to three minutes, we asked only five questions of the students. We began by asking why they did not respond to the NSSE, followed by four questions taken verbatim from the survey:

(a) How would you rate your relationships with faculty members? (Please use a scale from 1 (unavailable and unhelpful) to 7 (available and helpful) with four being right in the middle of the two extremes).

(b) Overall, how would you evaluate the quality of academic advising you have received at UMaine? (Please use excellent, good, fair, or poor).

(c) How would you evaluate your entire educational experience at UMaine? (Please use excellent, good, fair, or poor).

(d) If you could start over again, would you still go to UMaine? (Please use definitely yes, probably yes, probably no, or definitely no).

We chose these questions because they provide a view of a student's overall level of satisfaction and perceived relationships with faculty members—and without creating a complicated interaction between interviewer and interviewee. The questions are each self-contained (i.e., they do not require interviewees to use short-term memory to refer back to a previous question) and straightforward, and they offer simple response options. (The interview script is available upon request.)

In the analysis phase of the project, we compared the responses to the four questions above with the responses provided by UMaine NSSE participants. We used regression analysis to examine if there was a statistically significant difference between responders and the telephone interviewees in student-faculty relationship ratings. Because NSSE, in its standard reporting, takes into account level, gender, and enrollment status (either through a separation of results or a weighting procedure), we included these characteristics as control variables.

Specifically, we regressed the student-faculty rating on a nonresponder indicator variable (0 = NSSE responder, 1 = telephone interviewee), level (0 = freshman, 1 = senior), gender (0 = male, 1 = female), and enrollment status (0 = full-time, 1 = part-time).

To determine if responders and telephone interviewees differed in their overall satisfaction with their UMaine experience, we first collapsed categories to create dichotomous variables from the responses to each of the three overall satisfaction questions. We transformed the overall advising and overall educational experience responses to *good/excellent* or *fair/poor*, and the likelihood of returning to UMaine response to *definitely yes/probably yes* or *probably no/definitely no*. We then used logistic regression, a methodology which allows for a dichotomous dependent variable, to regress the transformed responses on, as before, the nonresponder indicator variable (0 = NSSE responder, 1 = telephone interviewee), level (0 = freshman, 1 = senior), gender (0 = male, 1 = female), and enrollment status (0 = full-time, 1 = part-time).

Results

Demographic Comparisons

Table 1 displays the overall and class-level response rates. Overall, approximately 21% of the 4,178 students who received invitations to complete the NSSE responded. There was a statistically significant difference in response rates between freshmen and seniors ($p < .01$): The freshman response rate was 19% compared with 24% among seniors. Such a difference is not unusual, however: NSSE reported response rates of approximately 25% and 28%, respectively (NSSE, 2011).

Table 1. *Response Rates by Class*

	Freshmen	Seniors	Total
Total Surveys Administered	2,057	2,121	4,178
Respondents	391	502	893
Non-Respondents	1,666	1,619	3,285
UMaine Response rates	19%	24%	21%
Overall NSSE response rates	25%	28%	27%

Note. Adapted from *UMaine Institutional Report 2011 (Respondent Characteristics),* National Survey of Student Engagement, 2011.

Tables 2 and 3 provide freshman and senior response rates broken down by gender, residency, enrollment status, transfer status, cumulative GPA, and college of major.[1] The tables show the results of the χ^2 tests of independence, which assess whether there were differences across groups in the proportion of students who responded to the survey, representing an over- or under-representation among respondents. For example, the results in Table 2 show a statistically significant difference related to gender. One can see that females had a higher response rate (26.5%) than males (15.3%) and, further, were more highly represented among the respondents (61.4% vs. 38.6%). This distribution of males and females is in contrast to that in the total student population (47.9% vs. 52.1%). If there was no relationship between gender and response, the split between females and males among the respondents would have been more reflective of that seen in the total population.

Among freshmen, a significantly higher proportion of females, first-time students, and students living on campus responded to the survey. As for seniors, more likely to respond were

[1] We created our dataset from the spring database, which only includes students who are registered for courses during the spring semester. Because the names supplied to NSSE were taken from the fall 2010 student database, some students who received surveys were not registered for classes in spring 2011. To restrict the sample to only students who were enrolled during the time of the survey administration, we excluded these students. This resulted in the exclusion of six freshmen and 13 seniors who responded to the NSSE survey.

females, full-time students, nonresidents, and students living on campus. Academic achievement, as measured by cumulative GPA, also was related to survey participation: For both freshmen and seniors, the response rate of students with GPAs above 3.5 was roughly double that of students with GPAs below 2.5. Across colleges,[2] there were differences in response rates among seniors but not among freshmen. The College of Natural Sciences, Forestry, and Agriculture enjoyed the highest senior response rate (32.5%), while the College of Business, Public Policy, and Health showed the lowest (18.6%).

[2] BPPH = College of Business, Public Policy, and Health, EHD = College of Education and Human Development, ENGR = College of Engineering, LAS = College of Liberal Arts and Sciences, NSFA = College of Natural Sciences, Forestry, and Agriculture. Further, DLL = Division of Lifelong Learning and EXPL = Explorations (a program for undecided students and/or students requiring additional academic preparation).

Table 2. *Demographic Comparisons: Freshmen*

Demographic Groups	UMaine Population	NSSE Respondents	Response Rate	% of Population	% of Respondents	χ^2 Test of Independence χ^2	*df*	*p*
Gender								
Females	896	237	26.5%	47.9%	61.4%			
Males	976	149	15.3%	52.1%	38.6%	35.7	1	< .01
Enrollment status								
Full-time	1,789	374	20.9%	95.6%	96.9%			
Part-time	83	12	14.5%	4.4%	3.1%	2.0	1	.16
Transfer status								
New	1,783	376	21.1%	95.4%	97.4%			
Transfer	86	10	11.6%	4.6%	2.6%	4.5	1	< .05
Residency								
Resident	1,501	306	20.4%	80.2%	79.3%			
Nonresident	371	80	21.6%	19.8%	20.7%	.3	1	.62
Living situation								
On-campus	1,513	332	21.9%	80.2%	86.0%			
Off-campus	359	54	15.4%	19.2%	14.0%	8.4	1	< .01
College								
BPPH	134	25	18.7%	7.2%	6.5%			
EHD	157	26	16.6%	8.4%	6.7%			
ENGR	325	61	18.8%	17.4%	15.8%			
NSFA	428	99	23.1%	22.9%	25.7%			
LAS	558	131	23.5%	29.8%	33.9%			
(EXPL)	223	33	14.8%	11.9%	8.6%			
(DLL)	47	11	23.4%	2.5%	2.9%	11.9	6	.07
GPA								
Below 2.5	685	108	15.8%	37.0%	28.1%			
2.5 - 2.99	407	64	15.7%	22.0%	16.7%			
3.0 - 3.49	408	93	22.8%	22.0%	24.2%			
3.5 or higher	352	119	33.8%	19.4%	31.7%	54.2	3	< .01

Table 3. *Demographic Comparisons: Seniors*

Demographic Groups	UMaine Population	NSSE Respondents	Response Rate	% of Population	% of Respondents	χ^2 Test of Independence χ^2	*df*	*p*
Gender								
Females	935	267	28.6%	46.5%	54.0%			
Males	1,075	227	21.1%	53.5%	46.0%	14.9	1	< .01
Enrollment status								
Full-time	1,570	408	26.0%	78.1%	82.6%			
Part-time	440	86	19.6%	21.9%	17.4%	7.7	1	< .01
Transfer status								
New	1,405	353	25.1%	70.2%	71.8%			
Transfer	596	139	23.3%	29.8%	28.2%	.7	1	.40
Residency								
Resident	1,758	420	23.9%	87.5%	85.0%			
Nonresident	252	74	29.4%	12.5%	15.0%	3.6	1	.06
Living situation								
On-campus	282	101	35.8%	14.0%	20.5%			
Off-campus	1,728	393	22.7%	86.0%	79.5%	26.1	1	< .01
College								
BPPH	247	46	18.6%	12.3%	9.3%			
EHD	225	56	24.9%	11.2%	11.3%			
ENGR	404	89	22.0%	20.1%	18.0%			
NSFA	502	163	32.5%	25.0%	33.0%			
LAS	608	138	22.7%	30.2%	28.0%			
(DLL)	24	2	8.3%	1.2%	0.4%	27.6	5	< .01
GPA								
Below 2.5	229	29	12.7%	11.4%	5.9%			
2.5 - 2.99	540	107	19.8%	26.9%	21.7%			
3.0 - 3.49	724	195	26.9%	36.1%	39.5%			
3.5 or higher	515	163	31.7%	26.4%	33.5%	40.2	3	< .01

In summary, the results of the χ^2 tests of independence show statistically significant differences across demographic groups. An overview of the groups with disproportionately high response rates appears in Table 4.

Table 4. *Summary of Demographic Comparisons*

More likely to respond among freshmen were:
• Females • New students (as opposed to transfers) • Students living on campus • Students with higher GPAs
More likely to respond among seniors were:
• Females • Full-time students • Students living on campus • Students in the college of NSFA • Students with higher GPAs

Although there were differences in response rates across demographic groups, such differences are only indicative of nonresponse bias if there were also discrepancies in responses to pertinent survey items. To identify whether such a threat exists, we conducted between-group comparisons using the demographic categories in Table 4 on the four items used in the nonresponder analysis (relationships with faculty members, quality of academic advising, perception of overall experience, and likelihood of attending again) and the five benchmark scores (Academic Challenge, Active and Collaborative Learning, Student-Faculty Interaction, Enriching Educational Experiences, Supportive Campus Environment). We used the χ^2 test of independence for the ordinal and nominal items and the independent sample *t* test or one-way analysis of variance for those with interval responses. The following is a summary of the statistically significant findings, with the complete results available upon request.

Among freshmen, there was a statistically significant difference associated with living situation: Freshmen living off campus had a more positive view of their interaction with faculty members (M = 35.8, SD = 20.5) than those living on campus (M = 29.3, SD = 16.8), $t(339) = 2.40$, $p < .05$. This difference corresponded to an effect size of approximately one third of a standard deviation ($d = .35$). Also, there was a significant effect associated with GPA and the Academic Challenge benchmark ($F(3, 329) = 2.81$, $p < .05$). However, follow-up pairwise comparisons were not statistically significant. Such a contradictory finding can be attributed to the conservative nature of the pairwise comparison test which, in order to reduce the probability of Type 1 error, corrects for the multiple comparisons being made.

The threat of nonresponse bias appears to be more pronounced among seniors, with discrepancies seen in relation to gender, college, and GPA. Males and females differed significantly in perceptions of their overall educational experience ($\chi^2(1, N = 423) = 5.61$, $p < .01$): 85% of females indicated their overall educational experience was *good* or *excellent* compared with 76% of males. In addition to being more satisfied with their overall experience, senior females also showed slightly higher Academic Challenge benchmark scores than males ($t(446) = 3.0$, $p < .01$), corresponding to an effect size of roughly one quarter of a standard deviation ($d = .28$).

Seniors of varying achievement levels differed in their perceptions of their relationships with faculty members ($F(2, 434) = 4.5$, $p < .01$) and the support provided by the campus environment ($F(3, 417) = 3.4$, $p < .05$). Tukey post-hoc comparisons indicate that students with GPAs of 3.5 or higher reported significantly more favorable sentiments regarding interactions with faculty (M = 5.5, SD = 1.2) than students with GPAs below 2.5 (M = 4.6, SD = 1.6), equaling an effect size of .30. Students with GPAs of 3.5 or higher also scored more highly on

the Supportive Campus Environment benchmark (M = 56.6, SD = 18.0) than students with GPAs of 2.5 or below (M = 43.6, SD = 23.5), corresponding to an effect size of .30.

Finally, seniors across colleges differed in their perceptions of the quality of academic advising they received ($\chi^2(4) = 11.5$, $p < .05$) and the support provided by the campus environment ($F(4, 416) = 5.0$, $p < .01$). ENGR seniors were the most positive in regard to academic advising (75% indicating *good* or *excellent*), while their LAS counterparts were the least positive (51% reporting *good* or *excellent*). Students in EHD were highest on the Supportive Campus Environment benchmark (M = 61.0, SD = 16.9), whereas students in NSFA (M = 51.4, SD = 18.6) and LAS (M = 51.8, SD = 18.4) were the lowest. The differences between the EHD students and those in NSFA and LAS correspond to effect sizes of .26 and .25, respectively.

Follow-up telephone interviews among nonresponders

Response rates. Table 5 displays a summary of the number of calls attempted, and the number and percentage of students who opted to participate in the telephone interviews. Approximately one third of the students with valid phone numbers for whom contact attempts were made participated in the interviews. Combined, almost 80% of the freshmen and seniors we ultimately were able to reach agreed to participate in the telephone interview. The telephone interviewees represent approximately 3% of the nonresponder population as a whole.

Table 5. *Response Rates for Follow-Up Analysis*

	Freshmen	Seniors
Total phone numbers attempted	193	175
Incorrect or disconnected numbers	41	31
Total valid numbers	152	144
Declined to be interviewed	11	15
Total telephone survey participants	50	50
Total not reached	91	79
Telephone survey participants as percentage of students reached	82%	77%
Telephone survey participants as percentage of valid numbers	33%	35%
Telephone survey participants as percentage of all nonresponders	3%	3%

Representativeness of telephone sample. Tables 6 and 7 report the demographic characteristics of the telephone interviewees compared with the characteristics of the nonresponder population as a whole. We used χ^2 goodness of fit tests to examine whether the differences between the two groups were statistically significant. The results show that the sample of interviewees was generally representative of the population of nonresponders. The one area where the nonresponder and telephone survey samples differed is in the proportion of freshmen in the various colleges: EXPL students were more highly represented among telephone interviewees than nonresponders (28% vs. 12.8%), while the opposite was true for BPPH (2% vs. 7.3%) and ENGR (12% vs. 17.8%) students.

Table 6. *Nonresponder Population vs. Telephone Survey Interviewees: Demographic Characteristics of Freshmen*

Demographic Groups	Nonresponder Population		Telephone Interviewees		χ^2 Goodness of Fit Test		
	n	%	*n*	%	χ^2	*df*	*p*
Gender							
Females	659	44.3%	23	46.0%			
Males	827	55.7%	27	54.0%	.06	1	.81
Enrollment Status							
Full-time	1,415	95.2%	47	94.0%			
Part-time	71	4.8%	3	6.0%	.16	1	.69
Transfer Status							
First-year	1,407	94.9%	49	98.0%			
Transfer	76	5.1%	1	2.0%	.99	1	.32
Residency							
In-state	1,195	80.4%	42	84.0%			
Out-of-state	291	19.6%	8	16.0%	.41	1	.52
Living Situation							
On-campus	1,181	79.5%	35	70.0%			
Off-campus	305	20.5%	15	30.0%	2.77	1	.10
College							
BPPH	109	7.3%	1	2.0%			
EHD	131	8.8%	3	6.0%			
ENGR	264	17.8%	6	12.0%			
NSFA	329	22.1%	11	22.0%			
LAS	427	28.7%	15	30.0%			
(EXPL)	190	12.8%	14	28.0%			
(DLL)	36	2.4%	0	0.0%	12.03	5	.03
GPA							
Below 2.5	566	39.4%	24	48.0%			
2.5 - 2.99	332	23.1%	14	28.0%			
3.0 - 3.49	307	21.3%	4	8.0%			
3.5 or higher	233	16.2%	8	16.0%	5.61	3	.13

Table 7. *Nonresponder Population vs. Telephone Survey Interviewees: Demographic Characteristics of Seniors*

	Nonresponder Population		Telephone Interviewees		χ^2 Goodness of Fit Test		
Demographic Groups	*n*	%	*n*	%	χ^2	*df*	*p*
Gender							
Females	668	44.1%	25	50.0%			
Males	848	55.9%	25	50.0%	.71	1	.40
Enrollment Status							
Full-time	1,162	76.6%	37	74.0%			
Part-time	354	23.4%	13	16.0%	.19	1	.66
Transfer Status							
First-year	1,052	69.7%	32	64.0%			
Transfer	457	30.3%	18	36.0%	.77	1	.38
Residency							
In-state	1,338	88.3%	42	84.0%			
Out-of-state	178	11.7%	8	16.0%	.90	1	.34
Living Situation							
On-campus	181	11.9%	5	10.0%			
Off-campus	1,335	88.1%	45	90.0%	.17	1	.68
College							
BPPH	201	13.3%	5	10.0%			
EHD	169	11.1%	5	10.0%			
ENGR	315	20.8%	8	16.0%			
NSFA	339	22.4%	13	26.0%			
LAS	470	31.0%	19	38.0%			
(DLL)	22	1.5%		0%	2.06	4	.73
GPA							
Below 2.5	181	12.3%	7	14.0%			
2.5 - 2.99	414	28.2%	13	26.0%			
3.0 - 3.49	520	34.4%	16	32.0%			
3.5 or higher	352	24.0%	14	28.0%	.70	3	.87

Reasons for nonresponse. When asked the reason for their non-response to the NSSE, those students who did remember receiving the survey—approximately two thirds— offered two primary reasons: they were too busy, or they just did not get to it or feel like it. There were only minimal differences between freshmen and seniors in response to this question. Table 8 shows the breakdown by class level.

Table 8. *Reasons for Nonresponse*

	Freshmen		Seniors	
	n	%	*n*	%
Don't remember the survey	19	38.0%	18	36.7%
Was too busy	14	28.0%	13	26.5%
Just did not get to it or feel like it	17	34.0%	16	32.7%
Other	0	0.0%	2	4.1%

Note. One senior did not provide a reason for nonresponse.

Relationships with faculty members. Overall, telephone interviewees indicated a slightly more favorable perception of their relationships with faculty members than did NSSE responders. Table 9 displays the mean ratings for responders and telephone interviewees.

Table 9. *Responders vs. Telephone Interviewees: Relationships with Faculty Members*

	Responders			Telephone Interviewees		
	n	*M*	*SD*	*n*	*M*	*SD*
Overall	761	5.2	1.4	96	5.5	1.1
Level						
Freshmen	323	5.0	1.4	48	5.3	1.2
Seniors	438	5.3	1.4	48	5.7	1
Gender						
Males	324	5.2	1.4	50	5.4	1.0
Females	437	5.1	1.4	46	5.6	1.2
Enrollment status						
Full-time	674	5.1	1.4	80	5.4	1.1
Part-time	87	5.3	1.3	16	5.9	1.1

Note. Four telephone interviewees did not respond to the question.

To determine if there was a statistically significant difference between responders and telephone interviewees, we regressed the faculty relationship rating on an indicator variable distinguishing telephone interviewees from the NSSE survey responders (0 = NSSE survey responders, 1 = telephone interviewees,). We included gender (0 = male, 1 = female), class level (0 = freshman, 1 = senior), and enrollment status (0 = full-time, 1 = part-time) as control variables. The results, which appear in Table 10, indicate that overall the model is statistically significant ($F(4,852)=3.27$, $p < .05$) and, further, there is a statistically significant difference in ratings between responders and telephone interviewees ($t = 2.31$, $p < .05$). Telephone interviewees reported ratings that were approximately one third of a point higher than responders (based on a seven-point scale). Class level also was a statistically significant predictor ($t = 2.43$, $p < .01$), with seniors showing ratings approximately one quarter of a point higher than those of freshmen. Neither gender nor enrollment status were statistically significant predictors.

Although the response and class-level indicator variables were both statistically significant, the magnitude of their effects is rather small. The R^2 shows that the model only explains approximately 1.5% of the variance in the ratings.

Table 10. *Predicting Student-Faculty Relationship Ratings*

	Coefficients		*Overall Model*	
	b	*t*	R^2	*F*
Constant	4.88**	27.60		
Telephone interviewee indicator	.34*	2.31		
Class level	.24**	2.43		
Gender	-.05	-.56		
Enrollment status	.09	.58	.015	3.27

* $p < .05$
** $p < .01$

Overall satisfaction. Tables 11-12 display student responses to questions about the quality of academic advising received and their overall educational experience; Table 13 shows the responses when asked whether or not they would still attend UMaine if they had the chance to start over. For ease of interpretation, we reduced the four-point scales to dichotomous variables: *excellent/good* vs. *fair/poor* for academic advising and overall experience, and *definitely yes/probably yes* vs. *probably no/definitely no* for likelihood of attending UMaine again.

There was little difference between responders and telephone interviewees in their perceptions of the overall quality of academic advising they received (75% vs. 72% for freshmen and 62% vs. 61% for seniors). In contrast, telephone interviewees were more likely than responders to rate their overall educational experience as good or excellent (96% vs. 85% for freshmen and 88% vs. 81% for seniors), and to indicate that they would attend UMaine again if they had the chance to start over (90% vs. 82% for freshmen and 96% vs. 75% for seniors).

Table 11. *Responders vs. Telephone Interviewees: Quality of Academic Advising*

	Responders				Telephone Interviewees			
	Poor/Fair		Good/Excellent		Poor/Fair		Good/Excellent	
	n	%	*n*	%	*n*	%	*n*	%
Overall	233	32.2%	490	67.8%	32	33.3%	64	66.7%
Level								
Freshmen	76	24.9%	229	75.1%	13	27.7%	34	72.3%
Seniors	157	37.6%	261	62.4%	19	38.8%	30	61.2%
Gender								
Males	99	32.5%	206	67.5%	20	40.0%	30	60.0%
Females	134	32.1%	284	67.9%	12	26.1%	34	73.9%
Enrollment status								
Full-time	207	32.4%	431	67.6%	28	35.0%	52	65.0%
Part-time	26	30.6%	59	69.4%	4	25.0%	12	75.0%

Note. Four telephone interviewees did not respond to the question.

Table 12. *Responders vs. Telephone Interviewees: Overall Educational Experience*

	Responders				Telephone Interviewees			
	Poor/Fair		Good/Excellent		Poor/Fair		Good/Excellent	
	n	%	*n*	%	*n*	%	*n*	%
Overall	124	17.2%	599	82.8%	8	8.1%	91	91.9%
Level								
Freshmen	45	14.8%	260	85.2%	2	4.1%	47	95.9%
Seniors	79	18.9%	339	81.1%	6	12.0%	44	88.0%
Gender								
Males	65	21.3%	240	78.7%	4	7.8%	47	92.2%
Females	59	14.1%	359	85.9%	4	8.3%	44	91.7%
Enrollment status								
Full-time	106	16.6%	532	83.4%	8	9.6%	75	90.4%
Part-time	18	21.2%	67	78.8%	0	0.0%	16	100.0%

Note. One telephone interviewee did not respond to the question.

Table 13. *Responders vs. Telephone Interviewees: Likelihood of Attending Again*

	Responders				Telephone Interviewees			
	Definitely no/Probably no		Probably yes/Definitely yes		Definitely no/Probably no		Probably yes/Definitely yes	
	n	%	*n*	%	*n*	%	*n*	%
Overall	160	22.1%	565	77.9%	7	7.1%	92	92.9%
Level								
Freshmen	54	17.6%	252	82.4%	5	10.2%	44	89.8%
Seniors	106	25.3%	313	74.7%	2	4.0%	48	96.0%
Gender								
Males	69	22.6%	236	77.4%	3	5.9%	48	94.1%
Females	91	21.7%	329	78.3%	4	8.3%	44	91.7%
Enrollment status								
Full-time	139	21.7%	501	78.3%	7	8.4%	76	91.6%
Part-time	21	24.7%	64	75.3%	0	0.0%	16	100.0%

Note. One telephone interviewee did not respond to the question.

To identify whether there was a statistically difference between responders and telephone interviewees in how they responded to these three questions, we conducted three separate logistic regression analyses in which we regressed each of the variables described in Tables 11-13 on the nonresponder indicator variable, gender, class level, and enrollment-status indicators. The results, which appear in Table 14, show no statistically significant difference between responders and telephone interviewees in their perceptions of academic advising. However, telephone interviewees were more likely than responders to indicate their overall educational experience was positive ($\chi^2 = 5.27$, $p < .05$) and that they would attend UMaine if given the chance to do it again ($\chi^2 = 10.02$, $p < .01$). Enrollment status was not a significant predictor in either model, but females were more likely than males to positively rate their overall educational experience ($\chi^2 = 5.26$, $p < .05$).

The odds ratio is helpful in interpreting the magnitude of these statistically significant effects. The odds of the telephone interviewees indicating they had a positive overall experience (i.e., *good* or *excellent*) was almost two and a half times that of the NSSE responders, holding level, gender, and enrollment status constant. The odds ratio was higher when it came to indicating whether they would return to UMaine if they had to do it again. Here, the odds of the telephone interviewees stating they would return if given another chance was roughly three and a half times that of the NSSE survey responders.

Table 14. *Predicting Overall Satisfaction Items*

	Overall Educational Experience			Likelihood of Attending Again		
	B	χ^2	Odds Ratio	*B*	χ^2	Odds Ratio
Constant	.78	3.23	2.18	-.22	0.25	0.80
Telephone interviewees	.88*	5.27	2.41	1.28**	10.02	3.59
Class level	.32	2.39	1.37	-.50*	7.03	0.61
Gender	.44*	5.26	1.55	.00	0.00	1.00
Enrollment status	-.05	0.03	0.95	.09	0.11	1.09

* $p < .05$
** $p < .01$

Discussion

Overview of Findings

The purpose of this study was to identify whether there is evidence of nonresponder bias in the NSSE data collected from UMaine students in spring 2011. Specifically, we addressed the questions of whether nonresponders and responders were demographically similar, why nonresponders did not participate in the survey, and how responders and nonresponders may have differed in their perceptions of UMaine.

The demographic analysis suggests NSSE respondents did differ on such key characteristics as gender, enrollment status, living situation, GPA, and college. Females, students living on campus, and students with higher GPAs were among the freshmen most likely to respond; and females, full-time students, those living on campus, majors in NSFA, and students with higher GPAs were among the seniors most likely to respond. Many of these differences are consistent with what has been reported in survey research specific to higher education. Males, part-time students, and students with lower GPAs typically are less likely to respond to surveys than their counterparts (Kuh et al., 2001; Porter & Umbach, 2006; Sax, Gilmartin & Bryant, 2003). Although the differences across colleges may be partially related to a general propensity among students in certain programs to respond to surveys, these differences

also likely reflect differences among colleges in efforts aimed at bolstering student response to the survey.

The results of the demographic analyses of responders reveals that, in some cases, groups more highly represented among responders may have had differing perceptions from those underrepresented. Among freshmen, specifically, there were differences associated with living situation; among seniors, there were differences among gender, GPA, and college. For example, seniors with GPAs above 3.5 (who were more highly represented among responders than those with lower GPAs) were also more positive about their interactions with faculty members and the level of support provided by the campus environment. Although there were discrepancies among disproportionately represented groups, which suggest some degree of nonresponse bias, the effects are rather small, with *d*'s ranging from .25 to .35.

As mentioned, one of the purposes of the follow-up analysis was to gain insight into why students did not respond to the survey. The results of the telephone interviews reveal two key findings in this regard. First, the majority of interviewees confirmed they did in fact receive the email invitation to participate in the survey. Second, interviewees offered a lack of time or motivation as their primary reasons for nonresponse. Although it is possible that either of these reasons could correlate with student engagement, there was no evidence to suggest that students consciously did not respond because of their dissatisfaction with UMaine.

With the follow-up analysis, we sought to identify whether responders and nonresponders differed in their perceptions of UMaine. To keep interviews simple and practical, we only included four overall measures of satisfaction rather than items specifically related to student behaviors. Although this limitation restricts one's ability to draw conclusions about levels of engagement, the questions do provide a view of students' level of satisfaction with their faculty interactions, the quality of advising they received, and the overall experience in general. These

are important considerations in their own right. And when combined with the question regarding a student's likelihood of returning to UMaine given a second chance, they can discriminate between students who had positive UMaine experiences and those having had less than positive experiences.

The results of the follow-up analysis suggest that, with the exception of academic advising, telephone interviewees had a more favorable view of their overall experience than responders. This should be interpreted with caution for two reasons. First, the use of telephone interviews is likely (and ironically) to introduce bias: Students may have a greater tendency to provide favorable answers on the phone than in a paper or online survey (Dillman et. al., 1996). In this respect, it is not surprising that the present results indicate more positive views among telephone interviewees. Second, the magnitude of some differences between the groups was quite small. For example, statistical significance notwithstanding, the telephone interviewees reported faculty interaction ratings that were only one third of a point (on a seven-point scale) higher than those of the responders.

These limitations, however, should not be interpreted as meaning that the telephone survey results are inflated and carry little import. If students were willing to express on the phone their discontent with their academic advising experience, they arguably would have been equally willing to disclose an unequivocal dissatisfaction with their UMaine experience. Erring on the side of caution, then, perhaps the best conclusion is by way of what these data do not suggest: They do not indicate that nonresponders are any less satisfied with their UMaine experience than responders.

Implications

Three major conclusions surface from our analyses: (a) there is evidence to suggest that NSSE responders are not demographically representative of the student population, (b)

nonresponse among students was primarily due to their lack of time or motivation, and (c) the evidence does not indicate that NSSE responders as a whole were more satisfied with their UMaine experience than nonresponders. Each of these conclusions has implications that will help UMaine better understand the NSSE results and their use.

Demographic differences. There were significant differences in response rates across demographic groups. Fortunately, the differences associated with gender and enrollment status are accounted for by a weighting procedure in NSSE's standard reporting. However, NSSE does not account for differences associated with living situation, GPA, and college of major. The discrepancies in the perceptions of students across groups within these demographic categories, although modest, are an issue that should be considered when interpreting the results. The campus-wide results will be more heavily impacted by the perceptions of overrepresented groups than those of the underrepresented, which could produce results that are more or less favorable than would have been found if respondents were demographically similar to the student population.

Reason for Response. The UMaine community can find comfort in the fact that there did not appear to be a problem with the process through which the survey invitations were sent out and, further, that students did not indicate their lack of response was due to strong negative feelings toward UMaine or the survey itself. In most cases, nonresponse was a function of a lack of time or general forgetfulness.

Nonresponse bias. A concern about the NSSE, and other surveys of a similar nature, is that students who respond may be more engaged in their college experience than those who did not (Kuh, 2003; Porter & Umbach, 2006; Porter & Whitcomb, 2005). The findings from the follow-up study fail to support such a concern, revealing no evidence to suggest responders were more satisfied with their experience than nonresponders. Further contradicting the presumption

that responders are more engaged, the demographic analyses revealed that students in underrepresented groups did not indicate consistently less favorable perceptions than students in overrepresented groups. For some groups on some items, it was the students more highly represented who were the most positive, but in other cases it was the underrepresented groups that responded more favorably.

Taken together, the findings from the demographic and follow-up analyses hold at least two implications regarding nonresponse bias. First, although there may be some degree of nonresponse bias, the results are not necessarily an inflated view of student perceptions. Second, any bias appears to be more a function of response rate differences across demographic groups than a fundamental difference between responders and nonresponders in general. The latter, coupled with the fact that the differences between overrepresented and underrepresented groups are small, suggests that disaggregating the results by demographic categories may help reduce the potential negative impact of nonresponse bias on the interpretation of the results. Had the follow-up study shown the perceptions of the nonresponders in general differed greatly from those of the responders, there would be little potential for minimizing the bias because the views of the nonresponders would remain unknown.

Concluding thoughts

The NSSE provides UMaine with meaningful feedback on student perceptions of their UMaine experience, and the results of the follow-up interviews suggest there is no reason to believe that responders as a whole are any more satisfied with their experience than nonresponders. However, some caution is advised when interpreting the results due to varying response rates and student perceptions across key demographic groups. Although the differences identified in this study are small and not likely to have a marked impact on the overall conclusions, they do point to the need for UMaine to be cautious when interpreting the campus wide results, and, in particular, to disaggregate the data beyond that provided in the standard NSSE reports. In our view, the UMaine results should be disaggregated by GPA, living situation, and college—the three areas where there were both response rate and perceptual differences in our data.

References

Banga, T.W., Pike, G., Hansen, M. (2009). The use of engagement data in accreditation planning and assessment. *New Directions for Institutional Research (141),* 21 – 34.

Bradburn, N.M. (1992). Presidential Address: A response to the non-response problem. *Public Opinion Quarterly*, 56, 391 – 398.

Chen, P.D. (2006). *Nonresponse effect in large-scale student surveys: Lessons learned from the National Survey of Student Engagement*. Paper presented at the NASPA International Assessment and Retention Conference, Phoenix, AZ.

Chen, P.D., Gonyea, R.M., Sarraf, S.A., BrckaLorenz, A., Korkmaz, A., Lambert, A.D., Shoup, R., Williams, J.M. (2009). Analyzing and interpreting NSSE data. *New Directions for Institutional Research (141),* 35 – 54.

Dey, E.L (1997). Working with low survey response rates: The efficacy of weighting adjustments. *Research in Higher Education, 18*(2), 215 – 227).

Dillman, D.A., Sangster, R.L., Tarnai, J., & Rockwood, T.H. (1996). Understanding differences in people's answers to telephone and mail surveys. *New Directions for Evaluation, 70*, 45 – 61.

Groves, R.M. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly, 70*(5), 646 – 675.

Hartman, B.W., Fuqua, D.R., & Jenkins, S.J. (1986). The problem of and remedies for nonresponse bias in educational surveys. *The Journal of Experimental Education, 54*(2), 85 – 90.

Kuh, G.D. (2003). *The National Survey of Student Engagement: Conceptual framework and overview of psychometric properties*. Bloomington, IN: Indiana Center for Postsecondary Research and Planning.

Kuh, G.D. (2009). The National Survey of Student Engagement: Conceptual and empirical foundations. *New Directions for Institutional Research (141)*, 5 – 20.

Kuh, G.D., Hayek, J.C., Carini, R.M., Ouimet, J.A., Gonyea, R.M. & Kennedy, J. (2001). *NSSE Technical and Norms Report*. Bloomington, IN: Indiana Center for Postsecondary Research and Planning.

McInnis, E.D. (2006). *Nonresponse bias in student assessment surveys: A comparison of respondents and non-respondents on the National Survey of Student Engagement in an independent comprehensive Catholic university*. Unpublished doctoral dissertation, Marywood University.

National Survey of Student Engagement. (2006). *Do students that respond to NSSE differ from those that choose not to respond to NSSE?* Retrieved from National Survey of Student Engagement website: http://nsse.iub.edu/pdf/psychometric_portfolio/OtherIndicators_NonresponseError_2005.pdf

National Survey of Student Engagement. (2008). *Nonresponse bias: Do previously engaged first-year students respond more?* Retrieved from National Survey of Student Engagement website: http://nsse.iub.edu/pdf/psychometric_portfolio/OtherIndicators_EngagedResponders.pdf

National Survey of Student Engagement. (2011). *Institutional Report for University of Maine*.

Pike, G. (2008). Using weighting adjustments to compensate for survey nonresponse. *Research in Higher Education, 49*(2), 153 – 171.

Porter, S.R., Rumann, C., & Pontius, J. (2011). The validity of student engagement survey questions: Can we accurately measure academic challenge? *New Directions in Institutional Research*, (150), 87 – 98.

Porter, S. R., & Whitcomb, M.E. (2005). Non-response in student surveys: The role of demographics, engagement, and personality. *Research in Higher Education, 46*(2), 127 – 152.

Porter, S. & Umbach, P. (2006). Student survey response rates across institutions: Why do they vary*? Research in Higher Education, 47*(2), 229 – 247.

Rogelberg, S. G. & Luong, A. (1998). Nonresponse to mailed surveys: A review and guide. *Current Directions in Psychological Science, 7*(2), 60 – 65.

Sax, L.J., Gilmartin, S., & Bryant, A.N. (2003). Assessing response rates and nonresponse bias in web and paper surveys. *Research in Higher Education, 44*(4), 409 – 432.