How Important are High Response Rates for College Surveys?

Kevin Fosnacht

Shimon Sarraf

Elijah Howe

Leah Peck

Indiana University Center for Postsecondary Research

Surveys play an important role in understanding the higher education landscape. About 60 percent of the published research in major higher education journals utilize survey data (Pike, 2007). Institutions also commonly use surveys to assess student outcomes and evaluate programs, instructors, and even cafeteria food. However, declining survey participation rates threaten this source of information and its perceived utility. Survey researchers across a number of social science disciplines in America and abroad have witnessed a gradual decrease in survey participation over time (National Research Council, 2013). Higher education researchers have not been immune from this trend as Dey (1997) long ago highlighted the steep decline in response rates in the American Council on Education and Cooperative Institutional Research Program follow-up surveys from 60 percent in the 1960s to 21 percent in 1991.

Survey researchers have long assumed that the best way to obtain unbiased estimates is to achieve a high response rate. For this reason, the literature on survey methods is rife with best practices and suggestions to improve survey response rates (e.g., American Association for Public Opinion Research, n.d.; Dillman, 2000; Heberlein & Baumgartner, 1978). These methods can be costly or require significant time or effort by survey researchers, and may be unfeasible for postsecondary institutions due to the increasing fiscal pressures placed upon them. However, many survey researchers have begun to question the widely held assumption that low response rates provide biased results (Curtin, Presser & Singer, 2000; Keeter, Miller, Kohut, Groves, & Presser, 2000; Groves, 2006; Massey & Tourangeau, 2013; Peytchev, 2013).

This study investigates this assumption for higher education assessment data. It utilizes data from hundreds of samples of first-year and senior students with relatively high response rates using a common assessment instrument with a standardized administration protocol. It investigates how population estimates would have changed if researchers put forth less effort when collecting data and achieved lower response rates and respondent counts. Due to the prevalence of survey data in higher education research and assessment efforts, it is imperative to better understand the relationship between response rates and data quality.

Literature Review

Survey nonresponse bias—the extent to which survey nonresponse leads to inaccurate population estimates—has received extensive attention in the survey research literature (e.g., Curtin, Presser, & Singer, 2000; Groves, 2006; Groves & Peytcheva, 2008; Rubin, 1976). Though variation exists with defining *nonresponse bias*, it is generally viewed as a function of the *response rate* and *nonresponse effects*, or how much responders and nonresponders differ on survey variables of interest (Keeter, et. al., 2000). In other words, low response rates may or may not lead to nonresponse bias because answers to survey items may not differ substantially between responders and nonresponders. The impact of nonresponse on an estimate depends upon the relationship between the outcome of interest and the decision to participate in the survey (Groves, 2006). Consequently, if the propensity to take a survey is not correlated with its content, the answers of responders and non-responders to a survey will not substantially differ. For these reasons, Massey and Tourangeau (2013) suggest that a high rate of nonresponse increases the *potential* for biased estimates, but does not necessarily bias an estimate. Peytchev (2013) goes farther and argues that the use of response rate as the singular measure of survey representativeness is flawed, as "it is nonresponse bias that is feared, not nonresponse itself" (p. 89).

Due to these insights, survey researchers have increasingly examined the impact of nonresponse on their survey estimates. Perneger, Chamot & Bovier (2005) assessed nonresponse bias by comparing outcomes between early-, late-, and non-responders. They found a modest difference in their estimated outcomes (less than .1 standard deviations) when comparing population estimates based on samples with only early responders (30% response rate) and the full sample (70% response rate). The authors concluded that while nonresponse bias did exist, greater survey participation "has only minimal influence on the conclusions of the survey" (p. 380). Similarly, using data from the Index of Consumer Sentiment (ICS), Curtin, Presser and Singer (2000) found no difference in their population estimates when comparing preliminary results based on response rates 5 to 50 percentage points lower than the final response rate. They created alternative estimates by excluding respondents that initially refused, required more than five recruitment calls, and required more than two recruitment calls. This analytical approach to assess population estimates under different response rate scenarios is generally referred to as a "level of effort" analysis (Olson, 2006), a term reflecting that a final response rate is somewhat artificial and dependent on when survey administrators stop contacting nonrespondents (or putting forth effort). Other health and psychology studies have come to similar conclusions based on results showing little variation under different response rate assumptions (Locker, 2006; Gerrits, van den Oord, & Voogt, 2001).

The results from these studies are not especially surprising given that other studies have found few differences between responders and nonresponders. Without a nonresponse effect, population estimates under different response rate scenarios should be highly correlated to estimates based on higher, final response rates. For instance, Mond et al (2004) determined in an eating disorder study that survey responses between first responders and those requiring several contacts did not differ. Additionally, a study of telephone survey responders found minimal differences between responses given by initial responders and those requiring several contacts to respond (Keeter, et. al, 2000).

In contrast, other researchers have found that increased efforts to collect survey data reduced nonresponse bias. One study, using household data from the German Panel Study, found that increased survey effort led to less nonresponse bias on a variety of individual characteristics (Kreuter, Muller & Trappmann, 2010). Unlike the other studies above, they had administrative information for the entire sample so an *absolute* estimate of nonresponse bias could be calculated. This differs from *relative* estimates of nonresponse bias obtained from studies that do not have a 100 percent response rate. However, this study evaluated non-response bias by examining individual's background characteristics rather than less-tangible measures like an individual's perceptions or satisfaction. Another study came to the same conclusion when examining patient satisfaction data on ratings of physicians and found substantial differences in their estimated outcomes (Mazor, Clauser, Field, Yood, & Gurwitz, 2002). Comparing the final population estimate to one of three simulated estimates, they found almost a full standard deviation difference, suggesting the potential for substantial nonresponse bias.

Others have found that nonresponse had varying effects on population estimates by comparing survey data on school characteristics to the same characteristics gathered from a secondary data source (Kano, Franke, Afifi, & Bourque, 2008). The authors found significant differences between responders and nonresponders for two (population density and enrollment in English Learner programs) of the seven variables studied. However, one of the variables, population density, was the only variable significantly related to survey response, thus demonstrating that biased estimates occur when response propensity is correlated with an outcome. This study also found that high-effort respondents did not significantly differ from low-effort respondents and nonrespondents using study variables.

A handful of higher education studies have focused on assessing survey nonresponse effect and bias. One study, based on about 600 first-year students enrolled in different classes assigned to different survey samples, did not find meaningful differences in students' perceptions of their academic environment when comparing estimates from administrations with response rates of 100 and 35 percent (Hutchison, Tollefson, & Wigington, 1987). Another series of studies conducted telephone interviews with randomly selected students who were asked to take the National Survey of Student Engagement (NSSE) multiple times, but failed to do so (Kuh, n.d.; Sarraf, 2005). These studies indicated that nonresponders responded differently to about half the tested survey items; however, they did not investigate the impact of nonresponse bias on institution-level population estimates. The authors cautioned that specific results indicating nonresponders to be more engaged may be the result of social desirability bias or telephone mode effects and not caused by true differences between responders and nonresponders. A third line of research examined the effectiveness of using survey weights to reduce nonresponse bias (Dey, 1997). It found that survey weights, derived from a regression predicting survey response, markedly improved population estimates and reduced nonresponse bias.

Several other higher education studies (Korkmaz & Gonyea, 2008; Porter & Umbach, 2006; Porter & Whitcomb, 2005; Sax, Gilmartin & Bryant, 2003; Sax, Gilmartin, Lee, Hagedorn, 2008) have focused on student and school characteristics associated with responding to surveys. However, these studies did not estimate how this might influence population estimates while taking into consideration response rates and nonresponse effect.

Theory

Survey nonresponse bias is a function of the nonresponse rate and the difference in means on an outcome between the respondents and nonrespondents. This relationship can be mathematically expressed as follows:

$$Bias_{NR} = Y_R - \bar{Y} = \pi_{NR} * (\bar{Y}_R - \bar{Y}_{NR})$$

Where, NR represents nonrespondent, R is respondent, and π_{NR} is the nonresponse rate. Unbiased estimates occur when the nonresponse rate or difference in means between responders and non-responders is zero. As researchers typically do not observe outcomes for nonresponders, they have traditionally emphasized reducing the nonresponse rate as much as possible to avoid obtaining unbiased estimates. Yet, unbiased estimates may also be obtained under conditions of high nonresponse when an outcome does not differ between respondents and nonrespondents.

Individuals will respond to a survey if they believe the benefits of participation will outweigh the costs. Leverage-saliency theory posits that when deciding to participate individuals assess a survey's features (e.g., topic, monetary incentive, organization) and their prominence in the request to participate (Groves, Singer, & Corning, 2000). Therefore, the effort exerted by a survey researcher plays a significant role in whether an individual participates in the survey, as incentives, customizing recruitment messages, and increasing the number of survey waves generally improve response rates (Goyder, 1982; Groves, Presser, & Dipko, 2004; Heberlein & Baumgartner, 1978). Thus, the response rate of a survey is a product of the characteristics of the potential respondents, the survey, and their interactions.

Research and theory on nonresponse generally overlooks the importance of surveyor effort. To demonstrate its importance, consider the effort expended by the Census Bureau when collecting data for the decennial census and by a grocery store who asks a customer to take a survey when paying for their items. In the former, the Census Bureau expends extraordinary effort when collecting data by administering multiple mailings, publicizing their efforts in the media, and making in person visits to collect data from nonresponders. In contrast, the store typically will ask the customer to respond once and may enter the respondent into a contest with a low probability of winning a monetary reward. Both of these surveys could easily change their characteristics by exerting more or less effort, which could result in a different response rate. Therefore, an individual's classification as a respondent or nonrespondent can vary, as their status may change due to different levels of effort exerted by the researcher.

Study Goals and Research Questions

This study seeks to investigate how survey population estimates vary under different response rate and respondent count assumptions from hundreds of college student survey administrations at a wide variety of North American colleges and universities. These findings can help initiate a robust discussion about survey data quality indicators and the role they play within the higher education community.

With these goals in mind, the following questions guided this study:

1) Do simulated *low response rate* survey estimates about college student engagement provide reliable information based on comparisons to actual high response rate estimates?
2) Do simulated *low respondent count* estimates provide reliable information based on comparisons to full sample estimates?
3) Do these results vary by survey administration size?

Methods

Data

We examined these research questions using data from the NSSE, one of the most widely used higher education assessment instruments. NSSE is annually administered to random or census samples of first-year and senior students using a standard protocol at hundreds of post-secondary institutions (National Survey of Student Engagement, 2012). This study's sample included data from online-only NSSE administrations between 2010 and 2012 that achieved a response rate greater than 50 percent and contained at least 20 respondents. 555 survey administrations at 307 institutions met these requirements. The distribution of response rates from the administrations included in this study by the number of students invited to take the survey and aggregated Carnegie Classification can be found in Table 1. Response rates varied between 50 and 100 percent with a median of 57 percent. The administrations meeting our inclusion criteria tended to ask less than 250 students to take the survey and occurred at institutions that only offered a bachelor's degree.

Our analyses focused on four NSSE measures: Level of Academic Challenge (LAC), Active and Collaborative Learning (ACL), Student-Faculty Interaction (SFI), and Supportive Campus Environment (SCE) benchmarks. These measures are composites of multiple survey items placed on a 100-point scale. Previous research has shown that the benchmarks produce dependable group means from samples as small as 50 students (Pike, 2012) and meet accepted standards of reliability and validity (National Survey of Student Engagement, 2013).

Analyses

Our data analysis was descriptive by nature. We first calculated means for each of the benchmarks at simulated response rates of 5, 10, 15, 20, 25, 30, and 35 percent for each survey administration included in the sample. We simulated the means by averaging the benchmark score of the initial respondents up to the response rate of interest. For example, if 100 students were invited to take the survey, the first five respondents, as measured by the time of survey submission, would be included in the simulated mean at a response rate of 5 percent. It should be noted that our data is *not* simulated or hypothetical; rather, we used observed data to simulate or re-estimate population means that would have been obtained under different response rate conditions.

For each of the benchmarks, we correlated the simulated means with the full sample mean. This approach is analogous to comparing the outcomes of the same survey administered with different levels of effort. The different levels of effort were hypothetical in this study, but could have been the product of factors such as a shorter field period, fewer reminder emails, generic invitations or reduced incentives. The correlations at these response rates were also calculated for very small ($20 < N < 250$), small ($250 \leq N < 500$), medium ($500 \leq N < 1{,}000$), and large ($N \geq 1{,}000$) administration sizes separately. We used a conservative correlation of .90 for evaluating reliability.

We repeated these analyses using respondent count as the level of effort indicator. The study examined survey estimates using simulated respondent counts of 10, 25, 50, 75, 100, 150, and 200 students. As with the response rate approach, we examined correlations between the full sample and simulated means across all institutions and by administration size.

Results

We initially investigated the correlations between the simulated benchmark estimates at different levels of effort and the full sample mean (see Table 2). At a simulated response rate of 5 percent, the correlations between the simulated estimate and the full sample estimate ranged from .64 to .76. After increasing the simulated response rate to 10 percent, all of the measures had correlations of .80 or higher. At 20 percent, the correlations for three of the benchmarks exceed .90, and the exception, SCE, nearly met this threshold at .89. Consequently, the full sample estimates were very similar to the simulated means at a 20 percent response rate. The correlations continued to rise along with the simulated response rate and approached 1.0 at a simulated rate of 35 percent.

Next, we examined the correlations by administration size. Stronger correlations were observed between the simulated and full sample means as the administration size increased. For administrations smaller than 250 students, the correlation between the simulated means at a five percent response rate and the full sample means ranged from .58 and .69. In contrast, we observed correlations between .93 and .97 for the same measures among administrations with at least 1,000 students. As with the overall results, the correlations between the simulated means and the full sample means rose along with the simulated response rate. The correlations for the very small administrations were greater than .90 for all measures when the mean was simulated to represent a 25 percent response rate. This bar was passed at simulated rates of 10 and 15 percent for the large and medium administration sizes, respectively.

After examining the results by response rate, we replicated the analyses by respondent counts (see Table 3). The correlations between a mean derived from just the first 10 respondents and the full sample ranged between .68 and .81 for the four measures studied. However, the correlations rose to between .86 and .92 after the simulated respondent count was increased to 25 students. The correlations between a simulated mean from the first 50 students and the full sample mean exceed .90 for all four measures when all of the administrations were included in the sample.

In contrast to the results by response rate, substantial differences between respondent count correlations were not observed by administration size. The correlations observed between a respondent count of 25 and the full sample means ranged between .86 to .93, .85 to .94, and .80 and .88 for ACL, SFI and SCE, respectively. The correlations between these measures were slightly less consistent for LAC and ranged from .74 to .90. Nearly all of the correlations between the full sample mean and the means derived from the first 50 respondents exceeded .90. The three exceptions surpassed this threshold when the respondent count was raised to 75 students.

Discussion

This study offers additional evidence that low response rate administrations can provide reliable survey estimates. Using over 500 first-year and senior student administrations from over 300 bachelor's degree-granting institutions, we found estimates for several measures of college student engagement to be reliable under low response rate conditions ranging from 5 to 25 percent, and as few as 25 to 75 respondents, based on a conservative reliability criteria ($r \geq .90$). These findings support the work of Hutchinson, Tollefson, & Wigington (1987) that show similar survey estimates of college student behaviors can be achieved based on a relatively low response rate administration. This study's results are not entirely surprising given the findings from NSSE nonresponder studies (Kuh, n.d.; Sarraf, 2005), as well as Pike's (2012) findings that NSSE benchmark scores based on 50 respondents provide dependable group means.

These results suggest that institutions and researchers examining college student behavior may not need to exert great effort maximizing response rates. Rather, the level of effort exerted by an institution can be contingent upon the size of the student population being examined. The results indicate that institutions with small enrollments need a relatively high response rate (20 to

25 percent) to be fairly confident in their survey estimates. In contrast, larger institutions can obtain reliable estimates with lower response rates. Regardless of administration size, a researcher's level of effort might be reduced, freeing time and monetary resources that could be better spent improving the survey instrument, analyzing the data or on other important projects.

The findings also suggest that researchers should pay more attention to minimizing sources of potential error, besides nonresponse, when evaluating data quality. We share Peytchev's (2013) concern that the overwhelming attention received by the response rate might distract from attending to other important types of survey error, such as measurement and sampling error. More emphasis should also be placed on investigating other data quality measures such as response differentiation, survey duration, and item nonresponse.

One important issue to review is whether the level of effort put forth by survey administrators should be guided by response rates or respondent counts. These results suggest that if you had to choose one, focusing on respondent counts would be wise, regardless of sample size. As stated previously, 25 to 75 respondents provided reliable estimates, whereas the response rate needed to achieve reliable estimates varied by administration size. Focusing solely on response rates may lead to confusion for survey administrators because of the varied response rates required across different administration sizes. However, response rates play a prominent role in data quality determinations by many constituents so they cannot be dismissed as irrelevant. As many know well, characterizing any individual survey administration as suffering from a low response rate will influence how results are received, regardless of how many individuals respond to a survey.

The vast majority of NSSE participating institutions conduct census administrations. Should this be standard practice? With college student survey burden being an issue that many campuses are struggling with, relying on random samples may be prudent for many institutions participating in NSSE, as well as other survey projects. Hypothetically, if your aim is to collect 50 respondents for a reliable estimate, and your population is 1,000, a reasonable approach would be to randomly sample 200 students, assuming a 25 percent response rate. The remaining 800 unsampled students could be used for other assessment projects, thus reducing overall survey burden and potentially increasing response rates for all surveys being administered on campus. This approach would require campus administrators to be more strategic with planning surveys for their campus, as well as requiring them to make accurate projections for ensuring a minimum respondent count. Survey administrators, such as NSSE, might also consider calculating for institutions an optimal sample size to yield a minimum number of respondents.

Despite the strong rationale for limiting the size of a survey administration, this approach holds some risks. Significantly fewer respondents will lead to less precise population estimates and a greater probability of making a Type II error when conducting statistical comparisons. Fewer respondents also mean institutions will have less data and power to investigate various student sub-groups on campus (e.g., academic major, ethnicity) and less confidence in these estimates. Before deciding to abandon census survey administrations, researchers should anticipate all possible impacts this would have on sub-group or future statistical analyses that the data may be used for.

A few study limitations should be noted before drawing any final conclusion. First, the study examined relative, not absolute, nonresponse bias. In other words, despite using relatively high response rate administrations in this study, knowing the true population statistic for all administrations could influence our results in some unanticipated way. Second, there is also the possibility that the NSSE schools that met our 50 percent response rate criteria from the NSSE

2010, 2011 and 2012 administrations are unique in a way that strongly influences our findings. For instance, the mean difference between early and late responders among schools with less than 50 percent response rates may be greater than the difference between these two groups at institutions within our study, thus resulting in lower reliability between simulated results and actual results.

Future investigations should help to shed light on identifying administrations that do not demonstrate reliable survey estimates with few respondents or low response rates. Nonresponders (or late responders) at some institutions may actually be very different than responders (or early responders), in which case exerting as much effort at boosting overall response rates and respondent counts would be warranted.

Conclusion

Survey administrators wanting to increase their response rate to an arbitrary number to satisfy external constituents should question whether their extra effort is warranted. This study did not find that a 5% response rate or even a 75% response rate provides unbiased population estimates under all circumstances, but rather that additional effort to move response rates marginally higher will frequently only shift survey results in trivial ways. Once survey administrators consider these results, we hope they will spend less time worrying about low response rates and more time evaluating and using the data they collect.

References

American Association for Public Opinion Research. (n.d.) *Best practices*. Retrieved from http://www.aapor.org/Best_Practices1.htm.

Curtin, R., Presser, S., & Singer, E. (2000). The effects of response rate changes on the index of consumer sentiment. *Public Opinion Quarterly, 64*, 413-428.

Dey, E. L. (1997). Working with low survey response rates: The efficacy of weighting adjustments. *Research in Higher Education*, 38(2), 215-227.

Dillman, D. A. (2000). *Mail and Internet surveys: The tailored design method.* 2nd ed. New York: J. Wiley.

Gerrits, M.H., van den Oord, E., & Voogt, R. (2001). An evaluation of nonresponse bias in peer, self, and teacher ratings of children's psychosocial adjustment. *Journal of Child Psychology and Psychiatry, 42,* 593-602.

Goyder, J. C. (1982). Further evidence on factors affecting response rates to mailed questionnaires. *American Sociological Review* 47(4), 550-553.

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

Groves, R. M. & Peytcheva, E. (2008). The impact of nonresponse rates on nonresponse bias: A meta-analysis. *Public Opinion Quarterly* 72(2): 167–89.

Groves, R. M., Presser, S., & Dipko, S. (2004). The role of topic interest in survey participation decisions. *Public Opinion Quarterly* 68(1):2–31.

Groves, R. M., Singer, E., & Corning, A. (2000). Leverage-saliency theory of survey participation: Description and an Illustration. *Public Opinion Quarterly.* 64(3), 299-308.

Heberlein, T. A. & Baumgartner, R. (1978). Factors affecting response rates to mailed questionnaires: A quantitative analysis of the published literature *American Sociological Review* 43(4), 447-462.

Hutchison, J., Tollefson, N., and Wigington, H. (1987). Response bias in college freshman's responses to mail surveys. *Research in Higher Education, 26*, 99–106.

Kano, M., Franke, T., Afifi, A.A., Bourque, L.B. (2008). Adequacy of reporting results of school surveys and nonresponse effects: A review of the literature and a case study. *Educational Researcher 37*(8), 480-490.

Keeter, S., Miller, C., Kohut, A., Groves, R.M., & Presser, S. (2000). Consequences of reducing nonresponse in a national telephone survey. *Public Opinion Quarterly, 64*, 125-148.

Korkmaz, A. & Gonyea, R. M. (2008). *The effect of precollege engagement on the likelihood of response to the National Survey of Student Engagement*. Paper presented at the annual meeting of the Association for Institutional Research, Seattle, WA.

Kreuter, F., Muller, G. & Trappmann, M. (2010). Nonresponse and measurement error in employment research. *Public Opinion Quarterly, 74*(5), 880-906.

Kuh, G. D. (n.d.). *The National Survey of Student Engagement: Conceptual framework and overview of psychometric properties*. Retrieved 3/12/2013 from nsse.iub.edu/pdf/conceptual_framework_2003.pdf

Locker, D. (1992). Effects of non-response on estimates derived from an oral health survey of older adults. *Community Dentistry and Oral Epidemiology, 21,* 108-113.

Massey, D.S., Tourangeau, R. (2013). Where do we go from here? Nonresponse and social measurement. *The ANNALS of the American Academy of Political and Social Science, 645*(2013 January), 222-236.

Mazor, K.M., Clauser, B.E., Field, T., Yood, R.A., & Gurwitz, J.H. (2002). A demonstration of the impact of response bias on patient satisfaction surveys. *Health Services Research, 37(5),* 1403-1417.

Mond, J., M., Rodgers, B., Hay, P.B., Owen, C., & Beumont, P. J. V. (2004). Nonresponse bias in a general population survey of eating-disordered behavior. *International Journal of Eating Disorders, 36,* 89-98.

National Research Council. (2013). *Nonresponse in Social Science Surveys: A Research Agenda*. Washington, DC: The National Academies Press.

National Survey of Student Engagement. (2012). Administration protocol and procedures Retrieved 4/19/2012, from http://nsse.iub.edu/_/?cid=266

National Survey of Student Engagement. (2013). NSSE's Commitment to Data Quality Retrieved 3/12/2013, from http://nsse.iub.edu/_/?cid=154

Olson, K. (2006). Survey participation, nonresponse bias, measurement error bias and total bias. *Public Opinion Quarterly* 70, 737-758.

Perneger, T. V., Chamot, E., & Bovier, P. A. (2005). Nonresponse bias in a survey of patient perceptions of hospital care. *Medical Care*, 43(4), 374-380.

Peytchev, A. (2013). Consequences of survey nonresponse. *The ANNALS of the American Academy of Political and Social Science, 645*(2013 January), 88-111.

Pike, G. R. (2007). Adjusting for nonresponse in surveys. In John C. Smart (Ed.), *Higher education: handbook of theory and research, vol. 22* (pp. 411-450). The Netherlands: Springer.

Pike, G. R. (2012). NSSE benchmarks and institutional outcomes: A note on the importance of considering the intended uses of a measure in validity studies. *Research in Higher Education, 54*(2), 149-170.

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

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

Rubin, D. B. (1976). Inference and missing data. *Biometrika* 63(3) 581-592.

Sarraf, S.A. (2005). *An evaluation of nonresponse effect in the National Survey of Student Engagement*. Unpublished internal technical report. Bloomington, IN: Center for Postsecondary Research, Indiana University.

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

Sax, L. J., Gilmartin, S. K., Lee, J., & Hagedorn, L. S. (2008). Using web surveys to reach community college students: An analysis of response rates and response bias. *Community College Journal of Research and Practice*, 32(9), 712-729.

Table 1
Response rate distribution of administrations included in the study by administration size[1] and Carnegie Classification

		Percentile						
	N	Min.	10	25	50	75	90	Max.
Administration Size								
Very Small (20≤N<250)	293	50	51	53	59	67	74	100
Small (250≤N<500)	168	50	51	53	55	60	68	76
Medium (500≤N<1,000)	74	50	51	52	55	60	64	76
Large (N≥1,000)	20	50	50	50	52	55	69	72
Carnegie Classification (aggregated)								
Baccalaureate	335	50	51	54	58	64	71	98
Master's	117	50	50	52	54	59	61	100
Doctoral	13	50	50	51	55	60	65	65
Other/Not Classified	90	50	51	52	59	68	78	94
Total	**555**	**50**	**51**	**53**	**57**	**63**	**71**	**100**

[1] Administration size is the number of students asked to take NSSE.

Table 2
Correlations between simulated response rate and full sample means by benchmark and administration size

	Simulated Response Rate						
	5	10	15	20	25	30	35
Level of Academic Challenge							
All administrations	.64	.80	.86	.91	.93	.95	.97
Very small	.61	.76	.83	.90	.92	.94	.96
Small	.69	.87	.91	.95	.95	.96	.98
Medium	.78	.91	.94	.95	.97	.98	.99
Large	.94	.98	.98	.99	.99	.99	.99
Active and Collaborative Learning							
All administrations	.76	.88	.93	.95	.96	.97	.98
Very small	.69	.82	.89	.93	.95	.96	.97
Small	.79	.90	.94	.95	.97	.98	.98
Medium	.93	.97	.98	.99	.99	.99	1.00
Large	.97	.99	.99	.99	1.00	1.00	1.00
Student-Faculty Interaction							
All administrations	.75	.87	.92	.95	.96	.97	.98
Very small	.68	.81	.89	.92	.95	.96	.97
Small	.82	.91	.95	.97	.97	.98	.99
Medium	.89	.96	.98	.99	.99	.99	.99
Large	.93	.98	.98	.99	.99	.99	.99
Supportive Campus Environment							
All administrations	.66	.80	.86	.89	.92	.95	.96
Very small	.58	.74	.81	.86	.90	.93	.95
Small	.79	.89	.94	.95	.95	.97	.98
Medium	.84	.90	.93	.94	.96	.97	.98
Large	.95	.97	.98	.98	.99	.99	1.00

Note: Very small = Less than 250 students sampled; Small = 250 through 499 students sampled; Medium = 500 through 999 students sampled; Large = 1,000 or more students sampled

Table 3
Correlation between simulated respondent count and full sample means by benchmark and administration size

	Simulated Respondent Count						
	10	25	50	75	100	150	200
Level of Academic Challenge							
All administrations	.68	.87	.94	.96	.97	.98	.99
N	*555*	*551*	*494*	*430*	*362*	*245*	*178*
Very small	.74	.90	.96	.98	.99	.99	---
N	*293*	*289*	*232*	*168*	*100*	*8*	*0*
Small	.58	.83	.92	.95	.97	.99	.99
N	*168*	*168*	*168*	*168*	*168*	*143*	*83*
Medium	.57	.74	.88	.91	.94	.97	.98
N	*74*	*74*	*74*	*74*	*74*	*74*	*74*
Large	.55	.76	.92	.94	.97	.97	.98
N	*20*	*20*	*20*	*20*	*20*	*20*	*20*
Active and Collaborative Learning							
All administrations	.81	.92	.96	.97	.98	.99	.99
N	*555*	*551*	*494*	*430*	*362*	*245*	*178*
Very small	.82	.93	.97	.98	.99	1.00	---
N	*293*	*289*	*232*	*168*	*100*	*8*	*0*
Small	.69	.86	.94	.96	.97	.99	1.00
N	*168*	*168*	*168*	*168*	*168*	*143*	*83*
Medium	.85	.92	.96	.98	.98	.99	1.00
N	*74*	*74*	*74*	*74*	*74*	*74*	*74*
Large	.84	.91	.93	.96	.97	.97	.98
N	*20*	*20*	*20*	*20*	*20*	*20*	*20*
Student-Faculty Interaction							
All administrations	.79	.92	.96	.98	.98	.99	.99
N	*555*	*551*	*494*	*430*	*362*	*245*	*178*
Very small	.82	.94	.97	.98	.99	1.00	---
N	*293*	*289*	*232*	*168*	*100*	*8*	*0*
Small	.70	.89	.95	.97	.98	.99	1.00
N	*168*	*168*	*168*	*168*	*168*	*143*	*83*
Medium	.77	.85	.95	.96	.97	.99	.99
N	*74*	*74*	*74*	*74*	*74*	*74*	*74*
Large	.78	.87	.90	.96	.95	.96	.97
N	*20*	*20*	*20*	*20*	*20*	*20*	*20*
Supportive Campus Environment							
All administrations	.70	.86	.93	.95	.96	.97	.98
N	*555*	*551*	*494*	*430*	*362*	*245*	*178*
Very small	.70	.88	.95	.98	.99	.96	---
N	*293*	*289*	*232*	*168*	*100*	*8*	*0*
Small	.71	.84	.93	.95	.97	.99	.99
N	*168*	*168*	*168*	*168*	*168*	*143*	*83*
Medium	.62	.80	.89	.91	.93	.96	.97
N	*74*	*74*	*74*	*74*	*74*	*74*	*74*
Large	.75	.86	.84	.91	.92	.95	.96
N	*20*	*20*	*20*	*20*	*20*	*20*	*20*

Note: Very small = Less than 250 students sampled; Small = 250 through 499 students sampled; Medium = 500 through 999 students sampled; Large = 1,000 or more students sampled