

RUNNING HEAD: The Institutional Quality Debate

**The Institutional Quality Debate:
U.S. News Quality Indicators and the National Survey of Student Engagement**

Paper presented at the 2005 meeting of the Association for the Study of Higher Education,
Philadelphia, Pennsylvania

Shimon Sarraf
George Kuh
John Hayek
Camille Kandiko
Ryan Padgett
Kim Harris

Center for Postsecondary Research
Indiana University
812-856-5824

Several initiatives are attempting to shift the conversation about collegiate quality away from resources and reputation toward more relevant indicators, such as student learning outcomes and graduation rates. The National Survey of Student Engagement (NSSE) is one such effort. NSSE was launched in 2000 and annually collects information directly from tens of thousands of students at several hundred colleges and universities. It measures the extent to which students engage in educational practices empirically linked to high levels of learning and development (Pascarella & Terenzini, 1991, 2005). Because the NSSE survey questions are based on previous research demonstrating links between participation in certain activities and desired benefits of attending college, student engagement results are thought to be reasonable proxies for institutional quality. To make the NSSE measure understandable and accessible to a wide range of groups on and off campuses, NSSE divided the survey items into five clusters or “benchmarks” of effective educational practice. They are level of academic challenge, student-faculty interaction, active and collaborative learning, enriching educational experience, and supportive campus environment.

Although NSSE and other instruments are being more widely used, their emergence has not, nor may it ever, replace the reputation and resource measures that have captured the public’s fancy. The best known are those that appear annually in *U.S. News and World Report* (referred to as *USNWR* for the remainder of this paper). The *USNWR* rankings are based on such indicators as selectivity, faculty salaries, and alumni giving. First published in 1983, the 3.5 million copies of *USNWR America’s Best Colleges* special edition reaches about 11 million people annually, generating more than \$5 million in revenue (Machung, 1998; Dichev, 2001).

However popular, the rankings have methodological flaws (Clarke, 2001; Hunter, 1995; McGuire, 1995; Stuart, 1995) which generally are ignored or dismissed by the public. For instance, after looking at the top 50 national universities according to *USNWR* Kuh and Pascarella (2004) concluded that “*USNWR* rankings of best colleges can largely be reproduced simply by knowing the average SAT/ACT scores of their students.” Equally problematic, “while higher education leaders are quick to criticize the annual rankings of American colleges and universities by *USNWR*, their institutions aggressively use the rankings to promote themselves in the race for prestige and visibility” (Machung, 1998, p.3). Many institutions outwardly refer to their rankings as a sign of quality in brochures, prospective student promotional mailers, and other public information sources (Greenberg, 2000; Hunter, 1995). Ultimately, rankings are an easy way to talk about reputation (Machung, 1998).

Another reason rankings continue to flourish is because the field of higher education has failed in presenting a persuasive case on behalf of alternative, more empirically compelling indicators. In addition, little information about these indicators is publicly accessible, making it almost impossible for people to compare colleges and universities on what are arguably more defensible measures of student and institutional performance (Greenberg, 2000).

Although NSSE was founded in part as an alternative to collegiate rankings in the assessment of institutional quality, relatively few studies have explored the relationships between these two sets of indicators. In 2001, using simple, institution-level partial correlations, researchers found few significant relationships between NSSE benchmarks and *USNWR* quality indicators (National

Survey of Student Engagement, 2001). Addressing problems associated with analyzing data at both the student and institutional level, Pike's (2004) study took a more sophisticated analytical approach using hierarchical linear modeling to test the relationships between NSSE benchmarks and *USNWR* rankings. Pike only found one significant relationship, which was between *USNWR*'s selectivity ranking and NSSE's enriching educational experience benchmark. However, Pike admittedly noted two major limitations in that he only used data from 14 schools, all relatively homogeneous public research universities.

Kuh and Pascarella (2004) found that the placement of the top 50 colleges ranked by *USNWR* could be essentially replicated by knowing one number – composite SAT or ACT scores of entering students. However, selectivity is unrelated to both student engagement and desired learning outcomes. They concluded, “institutional selectivity is a weak indicator of student exposure to good practices in undergraduate education” (p. 56). Thus, it appears that other dimensions of undergraduate learning are more important to student success than institutional reputation and selectivity (Pascarella, 2001).

Focusing on undergraduate student experiences in and out of the classroom, and the effort put forth, are better indicators of student outcomes (Hossler, 2000). NSSE was created in part to look at which factors of the undergraduate experience are positively associated with graduation and retention. Furthermore, there is no relationship between effective educational practice and alumni giving rates (Hossler, 2000). In regards to *USNWR*'s variable of expenditure per student, research has generally found no overall difference in spending between institutions that had both higher-than-predicted student engagement scores and graduation rates and peer schools (Gansemmer-Topf, Saunders, Schuh, and Shelley, 2004; NCHEMS, 2004).

Faculty resources, as measured by average class sizes, student faculty ratios, and percent of full-time faculty are positively related to successful undergraduate experiences through increased student-faculty contact (Astin, 1991; Chickering & Reisser, 1993). More personal interaction between students and faculty were associated with measures of education quality (Conrad & Conrad, 2001). However, the numbers may not matter as much as how the campus operates, as institutions with higher-than-predicted NSSE benchmarks scores had extremely varied student-faculty ratios (Kuh, Kinzie, Schuh, Whitt & Associates, 2005).

Purpose

This study explores the relationship between two commonly accepted measures of institutional quality – *USNWR* ranking indicators and NSSE's benchmarks of effective educational practice. More specifically, this study will explore the proportion of total variation in NSSE benchmark scores that can be explained at the student and institution level, what proportion can be explained by *USNWR* indicators of institutional quality, and, lastly, how individual *USNWR* variables such as peer academic reputation, retention and graduation performance, faculty resources, student selectivity, financial resources, and alumni giving relate to NSSE benchmarks.

Data

Sources. Four major data sources were used in this study. 2004 NSSE survey results provided all information for student background characteristics and constructing the five benchmark scores for each student. We obtained data for *America's Best Colleges 2004* from *USNWR*, which provided various institution-level indicators of quality. Since *USNWR* does not use a consistent set of variables to evaluate all types of institutions and certain information used to rank schools is not made public, it was necessary to supplement their dataset with IPEDS data. This lack of consistency also led to our using Barron's *Profile of American Colleges 2003* to provide a composite selectivity ranking that integrates high school rank, average standardized test scores, and acceptance rates (all of which are used by the *USNWR* ranking process).

Sample. All analyses for this paper used 62,594 senior NSSE respondents. Of this population, 65% were female, 9% underrepresented minorities, and 87% were enrolled full-time when they completed the survey in the spring 2004. These students came from 385 four-year colleges and universities that participated in NSSE 2004. On average, there were 164 student respondents from each institution. Institutions represented a diverse cross section of higher education with 21% being classified as Doctoral-Research, 46% Masters, 16% Baccalaureate-Liberal Arts, and 18% Baccalaureate-General. NSSE schools were excluded from this study either because they were not found in the *USNWR* dataset or had missing institution-level data.

Dependent Variables. Each of NSSE's five benchmarks represents aspects of student engagement in effective educational practices that have been associated with positive student outcomes (National Survey of Student Engagement, 2004):

- Level of Academic Challenge (LAC) measures the extent of challenging intellectual and creative work that is central to student learning and collegiate quality. Schools promote high levels of student achievement by setting high expectations for student performance.
- Active and Collaborative Learning (ACL) measures the degree that students are intensely involved in their education and asked to think about and apply what they are learning in different settings. Collaborating with others while solving problems or mastering difficult material prepares students to deal with life after college.
- Student-Faculty Interaction (SFI) attempts to quantify the frequency with which students learn firsthand from faculty members inside and outside the classroom. As a result of student-faculty interaction, teachers become role models, mentors, and guides for continuous, life-long learning.
- Enriching Educational Experiences (EEE) measures the extent that students participate in activities inside and outside the classroom that complement their academic programs. From experiencing diversity to taking advantage of internships and culminating senior experiences, certain experiences make learning more meaningful and, ultimately, more useful.

- Supportive Campus Environment (SCE) measures how students perceive their school’s commitment to their academic and social success. It also assesses how students feel about the quality of relationships they have with faculty, administrative personnel, and other students.

All the items contributing to the benchmarks were put on a 0 to 100 point scale and then averaged. (Appendix A contains more details in this regard.) All respondents answered at least three-fifths of the items on any given benchmark in order to receive a score. Once a raw score was calculated, we standardized each benchmark so that it had a mean of 0 and standard deviation of 1. Additional descriptive statistics about these variables and the independent variables described in the following section can be found in Table 1.

 Insert Table 1 about here

Independent Variables. Two sets of independent variables were used: student-level (level-1) and institution-level (level-2). Student-level variables included gender, ethnicity, first-generation college student status, student-athlete status, Greek-life status, full-time enrollment status, and academic major. Each of these student-level controls was represented by dichotomous variables or groups of dummy variables with an excluded group serving as reference.

Institution-level variables were solely represented by *USNWR* quality indicators or proxy in the case of selectivity. *USNWR* rankings were based on approximately 15 indicators that attempt to measure institutional quality, each of which is given a different weight in the final ranking calculation (Morse & Flanigan, 2004). These fifteen quality indicators can be grouped into seven areas: academic reputation peer scores, graduation and retention rates, faculty resources, student selectivity, financial resources, graduation rate performance (difference between expected and actual graduation rates), and alumni giving rate. Of the fifteen indicators, three were excluded from this analysis because complete information was not available: faculty compensation, percent faculty with terminal degree, and graduation rate performance. The twelve available indicators were either used by themselves or were combined together into composite scores. Not surprisingly, many of the indicators were highly correlated with each other and necessitated the creation of composite scores, or, in the case of student selectivity, borrowing another generally accepted measure like *Barron’s* selectivity rankings. That being said, if data could not be used directly from *USNWR*, every effort was made to substitute comparable information. Table 2 shows each *USNWR* variable included in the model, a short description of its contents, if and how it was transformed, and its source.

 Insert Table 2 about here

Modeling Approach

A multi-level modeling approach was employed using HLM 6.0. Multi-level modeling (or hierarchical linear modeling) has increasingly been used within the higher education research community to address questions where both individual and group-level variables relate to outcomes of interest. When presented with nested data (students grouped within institutions) and the outcome of interest is influenced by variables at both levels, OLS regression risks producing inaccurate coefficients, standard errors, and p-values for institution-level variables (Raudenbush & Bryk, 2002).

For each of the five NSSE benchmarks, four sequential HLM models were developed. First, a one-way ANOVA with random effects model (Model 1) determined how much variability exists within and between institutions with regard to NSSE benchmarks. Using these initial models, we calculated the intra-class correlation coefficient (ICC) to determine the proportion of benchmark variance at the institution level, and also confirmed that there was sufficient variability at the institution level to warrant the usage of institution-level predictors. Second, we computed a one-way ANCOVA with random effects model (Model 2) controlling for student-level variables such as race and gender. This model serves as the baseline to answer questions related to how much variance is explained at the institution level by using *USNWR* variables (controlling for student background characteristics). Lastly, two means-as-outcomes models incorporated *USNWR* variables with (Model 4) and without (Model 3) controls for institutional type. Model 3 provides both an estimate of how much between-institution and total variance in NSSE benchmarks is explained by *USNWR* variables after controlling for student characteristics. Model 4 uses the same approach as Model 3 but also controls for institutional type. Only parameter estimates for *USNWR* variables controlling for institutional type from Model 4 will be offered for interpretation.

Because the dependent measures (NSSE benchmark scores) were standardized, parameter estimates are equivalent to effect sizes where .10 is considered small, .30 moderate, and .50 large. Effect sizes can be interpreted as the proportion of a standard deviation change in a dependent variable given a one unit change in an independent variable. As we review our results, it is worth noting that some higher education researchers (Porter, 2005) argue even small effects can be important given the difficulty of affecting student behavior.

Results

Model 1. A relatively small amount of total variation in NSSE benchmarks can be explained at the institution level (see Table 3), although the amount is appreciably higher than earlier studies (Pike, 2004). Looking across all benchmarks, approximately 9% of the total variation in benchmark scores can be attributed to institutions. ICC statistics in higher education research are generally 10% or less so these results are not necessarily surprising or unique (Porter, 2005). Enriching Educational Experiences has the highest ICC (17%) and Active and Collaborative Learning the lowest (5%). Even though the majority of the variation can be attributed to students, the highly significant intercept variance components across all five benchmarks allow us to conclude that there is sufficient variation in average institutional benchmark scores to warrant continued modeling at the institution level using *USNWR* variables. Furthermore,

reliability statistics above .80 for each benchmark indicate that the average institutional benchmark score is a good estimate of the true institutional score.

 Insert Table 3 about here

Model 2. After controlling for various student background characteristics, Model 2 findings still indicated that there was sufficient variation between institutional mean benchmark scores to warrant continued modeling at level two. In addition, these models give us estimates of how much student-level variation is explained by our level-1 controls. On average, we explain about 4% of the variation at level-1. Enriching Educational Experiences had the most student-level variation explained (8%) and Supportive Campus Environment the least (1%). Because the focus of this study is on institution-level USNWR variables, parameter estimates for student-level variables are not presented here.

Model 3. By using results from Model 2 as a baseline, Model 3 results showed that including USNWR variables as predictors of average institutional benchmark scores had significant explanatory power. The portion of the between-institution variation explained by USNWR variables with all benchmarks except Active and Collaborative Learning exceeded 35%. Only 14% of the between-institution variation for Active and Collaborative Learning could be attributed to USNWR variables. Although variance explained statistics were unexpectedly high for four of the five benchmarks, the total variance (student and institution-level variation combined) explained with USNWR quality indicators was quite low. Only about 1 to 5% of the total variation in benchmark scores is associated with these variables. Variance components for the intercepts were still highly significant suggesting that there was additional variation between average benchmark scores that could be explained by other level-2 variables.

Model 4. Introducing institutional type into Model 3 increased the total variance explained between 5 and 11%, and helped to develop a better specified model so that USNWR parameter estimates would be less biased. Fifty-five percent of the between-institution variation in NSSE's Academic Challenge benchmark was accounted for after we added these additional controls; 25% for Active and Collaborative Learning; 53% for Student-Faculty Interaction; 52% for Enriching Educational Experiences; and 46% for Supportive Campus Environment.

In contrast to Pike (2004) who found almost no relationship between USNWR indicators and NSSE benchmarks, this study found a greater number of significant relationships, albeit with relatively small effect sizes (see Table 4). For instance, USNWR faculty resources showed a significant positive relationship with all five NSSE benchmarks. Student-Faculty Interaction and Supportive Campus Environment both show positive relationships to alumni giving rates, while Level of Academic Challenge and Enriching Educational Experiences positively relate to academic reputation peer scores. On the other hand, institutional selectivity shows no relationship to any of the NSSE benchmarks thus confirming the conclusions of recent studies (Kuh & Pascarella, 2004; Pascarella, Cruce, et al, 2005). Finally, though studies have shown increased educational expenditure per student to be unrelated to student outcomes, we have

found expenditures to be predictive of all but one NSSE benchmark, Active and Collaborative Learning.

 Insert Table 4 about here

Limitations

This study addressed at least one of the major limitations of previous attempts to look at the relationship between *USNWR* quality indicators and NSSE benchmarks—a small and homogenous group of institutions. Yet, three other caveats pertain to the models and results from this study.

First, not being able to obtain all *USNWR* information in the exact form that was used for calculating final rankings could lead to inaccurate results. For instance, average faculty compensation, an indicator that receives a relatively large weight by *USNWR* when estimating Faculty Resources, was excluded from our models because we were not able to adjust IPEDS salary information for cost of living. For other areas where we were able to obtain an estimate (e.g., educational expenditure per student), our variable value may have been different than that determined by *USNWR* because of the complex algorithms it uses to arrive at its numbers. To the extent measurement inaccuracies existed within our dataset, the results of this study are more or less reliable.

A second area of concern is how we specified our final set of models (Model 4). As with any OLS regression model, our multi-level approach is prone to biased coefficients because of unintentionally excluding key controls at both the student and institution-level. With the predominance of variance in NSSE benchmarks being explained at the student-level, and our inability to explain more than 10% of the variance on any given benchmark, a better specification of the level-1 model is essential for future research. Improving the degree of misspecification error at level-1 should ultimately lead to better estimates with institution level variables such as *USNWR* quality indicators. With this in mind, it is also quite possible that the significant relationships we see between NSSE outcomes and *USNWR* indicators could disappear once other non-*USNWR*, institution-level variables that are theoretically related to student engagement are included. Because our intention was to see if any relationship exists between *USNWR*'s view of institutional quality and NSSE outcomes, we did not set out to build a model that could explain all variance at the institution level.

Lastly, our models cannot be used to draw *definitive* causal conclusions between student engagement as represented by NSSE benchmarks and the variables that *USNWR* uses to rank schools.

Conclusions & Implications

The results of this study point to several conclusions. First, even though they differ dramatically in terms of their outward approach to institutional quality, both *USNWR* and NSSE indicators of

quality appear to share some common ground in terms of variance explained at the between-institution level. Similar to the conclusions drawn from many other higher education studies (Pascarella & Terenzini, 1991; 2005), the percent of variance in NSSE benchmarks explained by institutional level indicators was small, varying from only 5% to 17% compared to more robust within-institution differences. However, a major discovery was finding that between-institution variance accounted for by *USNWR* resource and reputation measures was surprisingly high for four out of five NSSE benchmarks, ranging from 36% to 49%. The one area of effective educational practice that *USNWR* indicators do not predict as well is Active and Collaborative Learning. This area of activity represents the nature of pedagogical approaches, so it's not surprising that *USNWR* measures do not capture this.

Second, as might be expected, institutions with greater faculty resources are more likely to show higher levels of performance on all engagement benchmarks. This institutional focus on the academic core appears to pay dividends in terms of engaging students in effective educational practices (Zemsky & Massy, 1995). The more students interact in smaller classroom settings and on campuses where the student-faculty ratio is smaller, the greater their likelihood of being challenged academically, experiencing active and collaborative learning, and feeling they are part of a supportive campus environment. A side effect of such engagement is that graduates of institutions with high overall engagement scores are more likely to contribute to their alma mater. This makes sense in that the stronger the quality of interpersonal relations between students, their peers, and their teachers, the more loyal they may be to the institution and more willing to support what they value.

In contrast, but consistent with other studies (e.g., Pascarella et al., in press), institutional selectivity and the NSSE benchmarks are independent of one another. This is additional evidence that the stress and anxiety of students getting admitted into more selective colleges and universities may be misguided to the extent highly selective schools are believed to superior in terms of engaging their students in educationally purposeful activities. Other studies also suggest that promising education practices are in place at diverse colleges and universities, many of which do not practice selective admissions (Kuh, Kinzie, Schuh, Whitt & Associates, 2005).

We cannot predict where the debate between resources, reputation, and student engagement will go next. What is clear, however, is that there are greater demands being placed on institutions of higher education from all directions, including accrediting agencies, state oversight commissions, governing boards, alumni, students, and the general public. Although NSSE's current participation agreement prevents the public dissemination of institutionally identifiable information, greater public demands for this information may necessitate greater dissemination of results. Further, the important findings about *USNWR* data elements are not easily available to the public since all of the measures are aggregated into a single, highly publicized index (Ehrenberg, 2003).

Kuh (1981) and Astin and Solomon (1981) argued more than 20 years ago that institutional quality was not unidimensional. This advice remains valid today. The changing demographics and societal expectations of higher education, along with its local and national impact, magnify the "high stakes" nature of looking at various methods to evaluate institutional quality and the college student experience (Kuh, 1995a; Kuh, Hu, & Vesper, 2000). Findings from the current

study, combined with the market success that both *USNWR* rankings and NSSE enjoy, appears to reinforce the need for a multi-dimensional look at institutional quality.

References

- Astin, A. W. (1991) *Assessment for excellence: The philosophy and practice of assessment and evaluation in higher education*. American Council on Education Series on Higher Education. Washington/New York: American Council on Education and Macmillan.
- Astin, A. W. (1993) *What matters in college? Four critical years revisited*. San Francisco: Jossey-Bass.
- Astin, A. W., & Solmon, L. C. (1981, October). *Are reputational ratings needed to measure quality? Change, 13(7)*, 14-19.
- Barron's Profiles of American Colleges (2003). *25th edition*. Hauppauge, New York: Barron's Educational Series, Inc.
- Clarke, M. (2001, April). *Quantifying quality: Technical issues surrounding U.S. News and World Report's rankings of U.S. colleges and graduate schools*. Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.
- Chickering, A. W., & Reisser, L. (1993). *Education and identity* (2nd ed.). San Francisco: Jossey-Bass.
- Conrad, M. J., & Conrad, M. A. (2001). Factors that predict academic reputation don't always predict desire to attend. *Journal of Marketing for Higher Education, 11(4)*, 1-18.
- Dichev, I. (2001). News or noise? Estimating the noise in the U.S. News university rankings. *Research in Higher Education, 42(3)*, 237-266.
- Ehrenberg, R. G. (2003, November). *Method or madness? Inside the "SNWR" college rankings*. Paper presented at the Wisconsin Center for the Advancement of Postsecondary Education Forum on The Use and Abuse of College Rankings, Madison, WI.
- Gansemer-Topf, A., Saunders, K., Schuh, J., Shelley, M. (2004). *A study of resource expenditures and allocation at DEEP colleges and universities: Is spending related to student engagement?* Retrieved September, 2005, from Indiana University, Center for Postsecondary Research web site: <http://nsse.iub.edu/html/research.htm>.
- Greenberg, M. (2000, June 16). America's colleges should rank themselves [The Chronicle Review]. *The Chronicle of Higher Education*.
- Hossler, D. (2000). The problem with college rankings. *About Campus, 5(1)*, 20-24.
- Hunter, B. (1995). College guidebooks: Background and development. In R. D. Walleri & M. K. Moss (Eds.), *Evaluating and responding to college guidebooks and rankings*. *New Direction for Institutional Research* (No. 88, pp. 5-12). San Francisco: Jossey-Bass.

- Kuh, G. D. (1981). *Indices of quality in the undergraduate experience*. AAHE-ERIC/Higher Education Research Report No. 4. Washington, DC.: ERIC Clearinghouse on Higher Education, and Washington, DC: American Association for Higher Education. (ERIC Document Reproduction Service No. ED213340)
- Kuh, G. D. (1995). Cultivating “high stakes” student culture research. *Research in Higher Education*, 36, 563-576.
- Kuh, G. D., Hu, S., & Vesper, N. (2000). They shall be known by what they do: An activities-based typology of college students. *Journal of College Student Development*, 41(2), 228-244.
- Kuh, G. D., Kinzie, J., Schuh, J. H., Whitt, E. J., & Associates. (2005). *Student success in college: Creating conditions that matter*. San Francisco: Jossey Bass.
- Kuh, G. D., & Pascarella, E. T. (2004). What does institutional selectivity tell us about educational quality? *Change*, 36(5), 52-58.
- Machung, A. (1998). Playing the ranking game. *Change*, (30)4, 12-16.
- McDonough, P. M., Antonio, A. L., Walpole, M., & Perez, L. X. (1998). College rankings: Democratized college knowledge for whom? *Research in Higher Education*, (39)5, 513-537.
- McGuire, M. D. (1995). Validity issues for reputational studies. In R. D. Walleri & M. K. Moss (Eds.), *Evaluating and responding to college guidebooks and rankings*. *New Directions for Institutional Research* (No. 88, pp. 45-59). San Francisco: Jossey-Bass.
- Morse, R., & Flanigan, S. (2004, May). *America’s best college rankings: A look back and a look ahead at plans for 2004*. Presentation at the annual national conference of the Association for Institutional Research, Boston, MA.
- National Survey of Student Engagement (2001). *Improving the college experience: National benchmarks for effective educational practice*. Bloomington, IN: Indiana University Center for Postsecondary Research.
- National Survey of Student Engagement (2004). *Student Engagement: Pathways to Collegiate Success*. Bloomington, IN: Indiana University Center for Postsecondary Research.
- National Center for Higher Education Management Systems [NCHEMS]. (2004). *Do DEEP institutions spend more or differently than their peers?* Retrieved September, 2005, from Indiana University, Center for Postsecondary Research web site: <http://nsse.iub.edu/html/research.htm>.
- Pace, C. R. (1980). Measuring the quality of student effort. *Current Issues in Higher Education*, 2, 10-16.

- Pascarella, E. T. (2001, May/June). Identifying excellence in undergraduate education: Are we even close? *Change*, 33, 19-23.
- Pascarella, E., Cruce, T., Umbach, P., Wolniak, G., Kuh, G., Carini, R., et al. (in press). Institutional selectivity and good practices in undergraduate education: How strong is the link? *The Journal of Higher Education*.
- Pascarella, E. T., & Terenzini, P. T., (1991). *How college affects students*. San Francisco: Jossey Bass.
- Pascarella, E. T., & Terenzini, P. T., (2005). *How college affects students: A third decade of research* (Vol. 2). San Francisco: Jossey Bass.
- Pike, G. R. (2004). Measuring quality: A comparison of U.S. News rankings and NSSE benchmarks. *Research in Higher Education*, 45(2), 193-208.
- Porter, S. (2005). *Institutional Structures and Student Engagement*. Manuscript submitted for publication.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Stuart, D. L. (1995). Reputational rankings: Background and development. In R. D. Walleri & M. K. Moss (Eds.), *Evaluating and responding to college guidebooks and rankings*. *New Direction for Institutional Research* (No. 88). San Francisco: Jossey-Bass.
- Zemsky, R. & Massy, W. F. (1995, November/December). Expanding perimeters, melting cores, and sticky functions: Toward an understanding of our current predicaments. *Change*, 41-49.

Table 1. Descriptive Statistics for U.S. News Indicators, NSSE Benchmarks, and Student Background Characteristics

Variable	Min	Max	Mean	St. Dev.
<i>Institution Level (n = 385)</i>				
U.S. News Quality Indicators				
Academic Reputation Peer Score	-2.43	3.14	0.00	1.00
Alumni Giving Rate	-1.60	3.82	0.00	1.00
Graduation & Retention Rate Composite	-2.67	2.51	0.00	1.00
Barron's Selectivity Rating (proxy for high school rank, average standardized test scores, and acceptance rate)	-2.19	2.76	0.00	1.00
Faculty Resources Composite	-2.39	2.91	0.00	1.00
Expenditure per student				
< \$10,000	0.00	1.00	0.13	0.34
\$20,000 to \$30,000	0.00	1.00	0.17	0.38
> \$30,000	0.00	1.00	0.04	0.20
Institutional Type				
Doctoral Research Universities-Extensive & Intensive	0.00	1.00	0.21	0.40
Masters Colleges and Universities I and II	0.00	1.00	0.46	0.50
Baccalaureate Colleges-General	0.00	1.00	0.18	0.38
<i>Student Level (n = 62,594)</i>				
NSSE Benchmark Scores				
Level of Academic Challenge (LAC)	-4.11	3.06	0.00	1.00
Active and Collaborative Learning (ACL)	-3.11	3.00	0.00	1.00
Student-Faculty Interaction (SFI)	-2.05	2.70	0.00	1.00
Enriching Educational Experiences (EEE)	-2.31	3.31	0.00	1.00
Supportive Campus Environment (SCE)	-3.24	2.28	0.00	1.00
Student Background Characteristics				
Student Athlete	0.00	1.00	0.07	0.26
Female Student	0.00	1.00	0.65	0.48
Underrepresented Minority Student	0.00	1.00	0.09	0.28
Multi-racial/other Student	0.00	1.00	0.05	0.22
Foreign Student	0.00	1.00	0.05	0.22
First Generation College Student	0.00	1.00	0.72	0.45
Greek Member	0.00	1.00	0.13	0.34
Fulltime Student	0.00	1.00	0.87	0.34
Live on or near campus	0.00	1.00	0.47	0.50
Double Major	0.00	1.00	0.13	0.34
Major: Biological Science	0.00	1.00	0.06	0.23
Major: Business	0.00	1.00	0.17	0.38
Major: Education	0.00	1.00	0.10	0.30
Major: Engineering	0.00	1.00	0.05	0.21
Major: Physical Science	0.00	1.00	0.03	0.16
Major: Professional	0.00	1.00	0.06	0.24
Major: Social Science	0.00	1.00	0.13	0.33
Major: Other	0.00	1.00	0.15	0.36

Table 2. Description of U.S. News Variables Used to Predict NSSE Benchmark Scores

Institution-Level Variables	Description	Model Transformation	Source
Academic Reputation Peer Score	Campus leaders asked to rate peer institutions' academic programs on a scale of 1 (marginal) to 5 (distinguished); all ratings averaged to produce final score	Standardized	2004 U.S. News
Alumni Giving Rate	Two-year average percentage of undergraduate alumni making a donation (00-01 & 01-02 AY)	Standardized	2004 U.S. News
Graduation & Retention Rate Composite	Combines 1997 freshman cohort 6-year graduation rate (80%) and average freshman retention rate for classes entering in 1999 to 2002 (20%)	Weighted t-score composite	2004 U.S. News 2003 IPEDS
Barron's Selectivity Rating	Combines standardized test scores, high school rank, and acceptance rates to measure institutional selectivity. A ranking of 1 to 11 is used (1 = "non-competitive"; 11 = "most competitive")	Standardized	2003 Barron's
Faculty Resources Composite	Combines percentage of classes with less than 20 students (60%), student-faculty ratio (20%), and percentage of full-time faculty (20%)	Weighted t-score composite	2004 U.S. News 2001 IPEDS
Expenditure per student	Total education expenses for the institution divided by total full-time equivalent enrollment was used to create 3 dummy variables: < \$10,000 \$20,001 to \$30,000 > \$30,000 <i>\$10,000 to \$20,000 (Reference Group)</i>	Dummy Coded 0 = no; 1 = yes	2002 IPEDS
Institutional Type	Using the Carnegie classification system, <i>USNWR</i> groups institutions into 4 categories based upon size and mission; 3 dummy variables are used: Doctoral Research Universities-Extensive & Intensive Masters Colleges and Universities I and II Baccalaureate Colleges-General <i>Baccalaureate Colleges-Liberal Arts (Reference Group)</i>	Dummy Coded 0 = no; 1 = yes	2004 U.S. News

Table 3. Variance Components and Variance Explained Statistics

		NSSE Benchmarks*				
		LAC	ACL	SFI	EEE	SCE
Model 1 ^a	Variance Component for Intercept	0.06 ***	0.05 ***	0.09 ***	0.17 ***	0.07 ***
	Intraclass Correlation (ICC)	6%	5%	9%	17%	7%
	Reliability estimate	0.87	0.85	0.91	0.96	0.90
Model 2 ^b	Variance Component for Intercept	0.05 ***	0.04 ***	0.05 ***	0.09 ***	0.07 ***
	Variance explained at level 1	3%	5%	4%	8%	1%
Model 3 ^c	Variance Component for Intercept	0.03 ***	0.03 ***	0.03 ***	0.05 ***	0.04 ***
	Variance explained at level 2	42%	14%	42%	49%	36%
	Total variance explained by US News variables	2%	1%	2%	5%	2%
Model 4 ^d	Variance Component for Intercept	0.02 ***	0.03 ***	0.02 ***	0.04 ***	0.04 ***
	Variance explained at level 2	53%	25%	53%	54%	45%
	Reliability estimate	0.76	0.78	0.75	0.86	0.82

* LAC: Level of Academic Challenge; ACL: Academic and Collaborative Learning; SFI: Student-Faculty Interaction; EEE: Enriching Educational Experiences; SCE: Supportive Campus Environment

^a Model 1: one-way ANOVA model where no student or institution-level predictors are used.

^b Model 2: one-way ANCOVA model where only student-level predictors are included (functions as the baseline model for models 3 and 4)

^c Model 3: means-as-outcomes model that adds US News variables at level-2 while controlling for student characteristics at level-1

^d Model 4: means-as-outcomes model using US News variables but also controlling for institutional type

+ p < .10; * p < .05; ** p < .01; *** p < .001

Table 4. Model 4 Coefficients for U.S. News Quality Indicators Predicting NSSE Benchmarks

	NSSE Benchmarks*				
	LAC	ACL	SFI	EEE	SCE
Intercept	-0.04	-0.42 ***	-0.06	-0.23 ***	-0.13 **
Academic Reputation Peer Score	0.03 *	0.00	0.01	0.07 ***	-0.02
Alumni Giving Rate	-0.01	-0.01	0.04 *	0.02	0.04 *
Graduation/Retention Composite	0.04 *	0.02	0.01	0.05 **	0.04 +
Barron's Selectivity Rating	0.01	-0.01	-0.01	0.01	0.02
Faculty Resources Composite	0.06 ***	0.05 ***	0.08 ***	0.08 ***	0.09 ***
Expenditure/student < \$10k	-0.03	0.00	0.00	0.01	0.06 +
Expenditure/student \$20 to \$30k	0.09 **	0.04	0.05 +	0.12 **	-0.06 +
Expenditure/student > \$30k	0.07	-0.04	-0.08 *	-0.08	-0.23 ***
<i>Reference Group: \$10-20k</i>					
National Universities	-0.30 ***	-0.22 ***	-0.26 ***	-0.28 ***	-0.21 ***
Masters	-0.18 ***	-0.06	-0.11 **	-0.25 ***	-0.01
Bac-Gen	-0.17 ***	-0.03	-0.06	-0.22 ***	0.07
<i>Reference Group: Bac-LA</i>					

* LAC: Level of Academic Challenge; ACL: Academic and Collaborative Learning; SFI: Student-Faculty Interaction; EEE: Enriching Educational Experiences; SCE: Supportive Campus Environment

Note: Student level characteristics controlled for include gender, ethnicity, parental education, greek, full-time status, academic major, and athletic status.

+ p < .10; * p < .05; ** p < .01; *** p < .001

Appendix A. Items Comprising NSSE Benchmarks, Response Sets, and Reliability Statistics

Benchmarks and Associated Survey Items	Response Set
Level of Academic Challenge ($\alpha = .753$)	
Hours per week preparing for class (studying, reading, writing, rehearsing, and other activities related to your academic program)	0, 1-5, 6-10, 11-15, 16-20, 21-25, 26-30, More than 30
Worked harder than you thought you could to meet an instructor's standards or expectations	Very often, often, sometimes, never
Number of assigned textbooks, books, or book-length packs of course readings during the current school year	None, 1-4, 5-10, 11-20, more than 20
Number of written papers or reports of 20 pages or more during the current school year	None, 1-4, 5-10, 11-20, more than 20
Number of written papers or reports between 5 and 19 pages during the current school year	None, 1-4, 5-10, 11-20, more than 20
Number of written papers or reports of fewer than 5 pages during the current school year	None, 1-4, 5-10, 11-20, more than 20
Coursework emphasizes: Analyzing the basic elements of an idea, experience, or theory	Very much, quite a bit, some, very little
Coursework emphasizes: Synthesizing and organizing ideas, information, or experiences into new, more complex interpretations and relationships	Very much, quite a bit, some, very little
Coursework emphasizes: Making judgments about the value of information, arguments, or methods	Very much, quite a bit, some, very little
Coursework emphasizes: Applying theories or concepts to practical problems or in new situations	Very much, quite a bit, some, very little
Campus environments emphasize: Spending significant amounts of time studying and on academic work	Very much, quite a bit, some, very little
Active and Collaborative Learning ($\alpha = .631$)	
Asked questions in class or contributed to class discussions	Very often, often, sometimes, never
Made a class presentation	Very often, often, sometimes, never
Worked with other students on projects during class	Very often, often, sometimes, never
Worked with classmates outside of class to prepare class assignments	Very often, often, sometimes, never
Tutored or taught other students (paid or voluntary)	Very often, often, sometimes, never
Participated in a community-based project as part of a regular course	Very often, often, sometimes, never
Discussed ideas from your readings or classes with others outside of class (students, family members, coworkers, etc.)	Very often, often, sometimes, never
Student Faculty Interaction ($\alpha = .760$)	
Discussed grades or assignments with an instructor	Very often, often, sometimes, never
Discussed ideas from your readings or classes with faculty members outside of class	Very often, often, sometimes, never
Received prompt feedback from faculty on your academic performance (written or oral)	Very often, often, sometimes, never
Talked about career plans with a faculty member or advisor	Very often, often, sometimes, never
Worked with faculty members on activities other than coursework (committees, orientation, student life activities, etc.)	Very often, often, sometimes, never
Work on a research project with a faculty member outside of course or program requirements	Done, plan to do, do not plan to do, have not decided
Enriching Educational Experiences ($\alpha = .640$)	
Practicum, internship, field experience, co-op experience or clinical assignment	Done, plan to do, do not plan to do, have not decided
Community service or volunteer work	Done, plan to do, do not plan to do, have not decided
Participate in a learning community or some other formal program where groups of students take two or more classes together	Done, plan to do, do not plan to do, have not decided
Foreign language coursework	Done, plan to do, do not plan to do, have not decided
Study abroad	Done, plan to do, do not plan to do, have not decided
Independent study or self-designed major	Done, plan to do, do not plan to do, have not decided
Culminating senior experience (comprehensive exam, capstone course, thesis, project, etc.)	Done, plan to do, do not plan to do, have not decided
Had serious conversations with students who are very different from you in terms of their religious beliefs, political opinions or personal values	Very often, often, sometimes, never
Had serious conversations with students of a different race or ethnicity than your own	Very often, often, sometimes, never
Institutional environment encourages contact among students from different economic, social, and racial or ethnic backgrounds	Very much, quite a bit, some, very little
Participate in co-curricular activities (organizations, campus publications, student government, social fraternity or sorority, intercollegiate or intermural sports, etc.)	None, 1-4, 5-10, 11-20, more than 20
Used an electronic medium (list-serv, chat group, Internet, etc. to discuss or complete an assignment)	Very often, often, sometimes, never
Supportive Campus Environment ($\alpha = .780$)	
Campus Environments Emphasize: Providing the support you need to help you succeed academically	Very much, quite a bit, some, very little
Campus Environments Emphasize: Helping you cope with your non-academic responsibilities (work, family, etc.)	Very much, quite a bit, some, very little
Campus Environments Emphasize: Providing the support you need to thrive socially	Very much, quite a bit, some, very little
Quality: Relationships with other students	1=Unfriendly, unsupportive, sense of alienation; 7=friendly, supportive, sense of belonging
Quality: Relationships with faculty members	1=Unavailable, unhelpful, unsympathetic; 7=Available, helpful, sympathetic