If You Build It, They Will Assess It (or, An Open Letter to George Kuh, with Love and Respect)

Michael A. Olivas

PREFACE

If You Build It, They Will Assess It (or, An Open Letter to George Kuh, with Love and Respect)

Michael A. Olivas

For large-scale college student assessment scholars, this is the best of times and the worst of times. It is the best, as they are the leading edge of a major paradigm shift in the polity, with more states wanting or needing to assess their schoolchildren through various No Child Left Behind (NCLB) requirements, and more college trustees requiring the same for college student counterparts. No good deed goes un-assessed these days. This has meant that there are many sources of support and development for the scholarship of evaluating these evaluations, and virtually no group does this as well as the Indiana University scholars loosely or directly affiliated with the various iterations and versions of Surveys of Student Engagement (SSE). This remarkable enterprise is a small empire (I write this with admiration and not a little envy, and not with Marxian overtones), housed in the Center for Postsecondary Research (CPR) in Indiana University's School of Education. The SSE is a tool widely employed by a rising number of institutions concerned with better understanding their students, particularly measuring the three leading strands evident in student development theory: involvement (most notably the construct pioneered by Alexander Astin [1993; 1999]), integration (largely defined by the work of Vincent Tinto [1993]), and

MICHAEL A. OLIVAS holds the William B. Bates Distinguished Chair of Law and is Director of the Institute of Higher Education Law and Governance, University of Houston Law Center. Address queries to him at 100 Law Center, Houston, Texas 77204; telephone: (713) 743–2078; fax: 713–743–2085; email: molivas@uh.edu.

most recently and successfully, engagement (based on the work of George Kuh and his collaborators, who in turn burrowed deeply into the work of Arthur Chickering and Linda Reisser [1993]). Each of these movements has its adherents and critics, but the rise of the SSE surveys was spawned in the tidal wave of the NCLB-related ethos, where assessments matter at all levels, including structural didactic shifts, such as the widespread use of instructional technology, asynchronous learning, and web-based teaching. Adding to this movement is the restructuring of the professoriate to make it more contingent and "efficient" and the astonishing growth of proprietary colleges that use such data for marketing purposes. Several recent federal government studies have shown the extent to which this sector aggressively engages in deceptive marketing and recruiting practices (U.S. GAO, 2010; U.S. Senate, 2010). As such, the rise of these assessments is equally, in my view, the worst of times.

The NSSE project situates itself within this larger stream, and its website notes: "Institutions use their data to identify aspects of the undergraduate experience inside and outside the classroom that can be improved through changes in policies and practices more consistent with good practices in undergraduate education. This information is also used by prospective college students, their parents, college counselors, academic advisers, institutional research officers, and researchers in learning more about how students spend their time at different colleges and universities and what they gain from their experiences" (NSSE, 2011a). It touts having more than 1,400 institutions in North America as NSSE users in the decade of its existence (it was founded in 2000); and, like Google and other e-tools, it both gathers and churns data so that its numbers grow, and repeat users find themselves a part of the larger project, providing extraordinary data in a snowballing fashion (Kuh, 2000a).

They have sliced the basic SSE approach finer and finer, extending it downstream (to community colleges and two-year institutions) and upstream (to law students). My recent analysis of the SSE enterprise shows the following outcomes projects, revealing how entrepreneurial and focused the Indiana group has become:

- Beginning College Survey of Student Engagement
- Classroom Survey of Student Engagement
- College Student Experiences Questionnaire Assessment Program
- Community College Survey of Student Engagement
- Faculty Survey of Student Engagement
- High School Survey of Student Engagement
- Law School Survey of Student Engagement
- National Institute for Learning Outcomes Assessment
- NSSE Institute for Effective Educational Practice
- Project on Academic Success
- Strategic National Arts Alumni Project

If imitation is the sincerest form of flattery, surely critique is its equal. And the papers assembled in this special issue of *The Review of Higher Education* are at the front end of this sincere critique, raising genuine concerns and taking the various SSE pieces on in a way that reveals both shock and awe, both respect and critical judgment. Because each article is so well contained and eloquent, I will summarize only the basics and conclude this concordance by suggesting how seriously the SSE folks should take the critiques. To have critics look your work over as carefully as has been the case(s) here, set out thoughtful corrections or questions, and recommend changes that will improve the overall project are gold; deep listening will not only improve the product but will allow it to become even more widespread and useful than it is at present. Perhaps making a virtue of necessity, Professor Kuh has allowed that the instruments and measures are able to be improved; he has conceded that "there are things about NSSE that aren't perfect in terms of its measures" (qtd. in Wolf-Wendel, Ward, & Kinzie, 2009, p. 421) and yet has also acknowledged, "If we were doing it again or we weren't worried about people using it over time we would change things now. We would add or subtract things—but when you are in year eight people don't want you to change things" (p. 421). Efficacy and consistency aside, the Kuh approach has changed the dialogue and discourse, just as Astin and Tinto did earlier. His desire, largely achieved, has been to "change the way people talk and think and act about what matters to collegiate quality and student learning" (p. 421). However, as these papers collectively and individually reveal, change for the sake of change is no better than the false god that was worshipped earlier. At the least, assessment and evaluation should do no harm.

In "Theoretical Foundations and a Research Agenda to Validate Measures of Intercultural Effort," Alicia C. Dowd, Misty Sawatzy, and Randi Korn (2011) have particular concerns about NSSE's and especially the Community College Survey of Student Engagement's "construct underrepresentation," borrowing from Goodwin and Leech's 2003 work, which they ground in the construct validity of "student effort" and a critical analysis of "intercultural effort":

> In order to bridge the theoretical foundations of the concept of student effort, which is based on human capital theory, and the concept of intercultural effort, which is based on cultural and critical theories, in the conceptual framework section, we first describe the historical tension between economic theories of capital (human, social, and cultural) and the cultural theories that critique them. Drawing on behavioral economics, we then model discrimination as a constraint on the postsecondary education (human capital) investments of students who encounter discrimination, adopting this approach to provide clarity to the measurement task that must be taken up. (p. 21)

In considerable detail, and in very clear terms, they question the validity of the various NSSE benchmarks, particularly the survey's lack of theoretical jus-

tification (in quantitative argot, the misspecification of theoretical models), its overarching content domain, and its vague justification for item inclusion:

> These arguments concerning the validity evidence of the NSSE are also applicable to the CCSSE, because the five CCSSE benchmarks are derived from and closely related to NSSE's five benchmarks. Both instruments aim to measure the constructs of academic challenge, active and collaborative learning, and student-faculty interaction. Whereas the NSSE measures supportive campus environments, the CCSSE measures supports for learners. "Student effort" is one of the CCSSE's five benchmarks. The NSSE, in contrast, does not label any set of its indicators as "student effort." Instead it includes indicators similar to those included in the CCSSE construct of effort in the "level of academic challenge" and "active and collaborative learning" scales, such as frequency of hours studied, reports written, or hours spent tutoring others [citations omitted].
>
> The theoretical construct of "student effort" is foundational to the indicators and response scales of the engagement surveys, which measure frequency of effort in various academic, collegiate, and interpersonal activities. Yet, according to Tanaka (2002), the measures omit the important construct of intercultural effort. Therefore, in order for the engagement benchmarks to be considered an inclusive "student effort" instrument, they would need to measure intercultural effort. An inclusive instrument would have the ability to measure all aspects of "student effort" including latent qualities such as effort to counter the effect of marginalizing experiences within the educational environment. The practical implications of creating a more inclusive instrument would be an improved reading of institutional effectiveness in evoking "student effort." (p. 22)

Dowd, Sawatzy, and Korn then employ this approach in a very careful march through the literature, which is extremely well-synthesized and useful. It is also exhibit No. 1 for the proposition that women scholars and people of color scholars are writing this narrative, particularly Latinos, African Americans, and Asians. Anyone attending ASHE today will not recognize it as your father's ASHE; and the scholarly efforts in critical theory, critical race theory, queer theory, feminist theories, and other forms of these genres are cascading into the literature as a result. Chicana scholar Sylvia Hurtado's taking over the UCLA Higher Education Research Institute and its venerable CIRP data collection from Sandy Astin upon his retirement is but one example of how the tides have shifted—generationally, substantively, and symbolically. The attention paid to campus climate measures, the involvement of critical theorists in the *Gratz* and *Grutter* versus University of Michigan affirmative action cases (whose work convinced the U.S. Supreme Court that diversity was a legitimate and constitutional criterion in college admissions), and the rising tide of books on these subjects—all augur a different discourse than that which arose from earlier work on predominantly White student bodies.

One of the more remarkable differences between scholars in this large field today and when I was starting out in the late 1970s is how the work today generates more light than it does heat. When we were the shock forces of the time, we cursed the vast darkness and generated a great deal of heat. Dowd, Sawatzky, and Korn are more grounded than we were and less reliant on tropes and parables. After their thorough review, here is their call to action:

> A comprehensive research agenda is needed to develop, pilot test, and validate new instruments that include measures of intercultural effort (Museus & Maramba, 2011; Nuñez, 2009; Tanaka, 2002). We contribute to that agenda through this review and analysis. (p. 21)

Not exactly storming the barricades, but certainly more likely to be taken seriously and to effect change that will survive and take root. But only if the NSSE people are listening.

Stephen R. Porter's (2009) estimable "Do College Student Surveys Have Any Validity?" (reprinted in this issue) takes the Dowd, Sawatzy, and Korn concerns about the limitations in the various SSE instruments to the next level, suggesting that the "typical college student survey question has minimal validity" and arguing:

> Our field requires an ambitious research program to reestablish the foundation of quantitative research on students. Our surveys lack validity because (a) they assume that college students can easily report information about their behaviors and attitudes, when the standard model of human cognition and survey response clearly suggests they cannot, (b) existing research using college students suggests they have problems correctly answering even simple questions about factual information, and (c) much of the evidence that higher education scholars cite as evidence of validity and reliability actually demonstrates the opposite. (pp. 45–46)

Yikes! And then he lays out his critique in a model of lucidity, thoroughness, and fairness—and not just NSSE, the biggest game. He dubs it the "preeminent survey of college students" (p. 46) and then skewers it, but he concedes offhandedly that his logic would apply to the CIRP data as well—take that, Professor Hurtado!

He takes a different tack on NSSE, arguing that it does not square up with the important literature on or experience with self reports, which he identifies as a major failure of most assessment instruments, especially those that are not time-use panels or time-use diaries. (Two side notes here: I had never heard of the Berea Panel Study of college students [Stinebrickner & Stinebrickner, 2004]; so when he mentioned it, I ran to Google-Scholar, poked around in that wonderful database, and found a fascinating small library of studies that have many of the desiderata that all the authors in this issue called for; second, he cited an old piece on self-reports I'd published more than 25 years

ago and have always loved [Olivas, 1986].) It had never occurred to me, as a regular reader of these assessment studies, that most fail on the grounds of poor self-reporting. Porter sensibly suggests:

> We as a field should abandon questions that appear to contradict theory and research on human cognition, which research shows are not accurate representations of behavior. The most obvious set of questions here is the self-reported learning gains that are used in so many college student surveys. It is clear that the majority of people cannot accurately report their learning gains, a finding confirmed by recent research. Questions asking students to report the number of hours spent on various activities should also be abandoned, unless the reference period is quite short. Most importantly, we should recognize that, for the majority of topics we are interested in, students can accurately report only on the previous week, and not the current semester or academic year. (pp. 72–73; internal citations omitted)

Gawd, could it actually be that easy? He goes on to note:

> In terms of evaluating surveys, we should understand that it is fairly easy to find small correlations between variables and that correlations can be misleading without any additional analyses. We should also establish criteria for judging before beginning validation research; for example, what do we mean by "highly correlated"? Much of the current validation research in higher education appears to take what I think of as the "greater than zero" approach—that is, if a correlation, standardized regression coefficient or reliability measure is greater than zero, then all is well. Clearly we as a field need to establish some commonly accepted minimum levels for judging relationships. (p. 73)

From his lips to God's ears. In my own work over the years, I am embarrassed to think how many times I have deduced that one variable was "highly correlated" with another; while I can remember every editorial slight over the years, I cannot recall a single time where a referee or editor rebuked me for the careless language and intellectual laziness this term reveals and hides. Indeed, other papers in this issue use the term as if it were self-evident. Porter also uses the secret vice of self-servingness as a counterpoint, and, boy, does it ring true—and do not bother to ask to see my self-reported demographics on my driver's license, at least without a subpoena:

> Researchers should also seek stronger evidence of validity than we have currently done in the past. For a given survey, it is relatively easy to look for convergent or divergent validity within the data. Consider again the example of asking students their height and weight in a survey. Higher education researchers might validate their findings by calculating the correlation between the reported number of desserts consumed per month and self-reported weight, or whether students on the basketball team report being taller than students on the chess team. Finding a positive correlation or difference, they would

conclude that the questions are valid, missing entirely the fact that students overreport their height and underreport their weight and that these errors are correlated with student characteristics. (p. 73)

But he saves his biggest IED in the road to Baghdad for this simple burden of proof, or of persuasion: "Finally, and most importantly, the tacit agreement in postsecondary research seems to be that validity is assumed until proven otherwise. Instead, we must establish standards such that a lack of validity is assumed until proven otherwise" (p. 73). If implemented and observed widely, this admonition would be the Miranda warning of educational research; and even though the U.S. Supreme Court has trimmed back Miranda (Guerra Thompson, 2006), it should guide practice and set moral standards for the practice and praxis of social science scholarship.

In the third paper in this special issue, "How Sound Is NSSE? Investigating the Psychometric Properties of NSSE at a Public, Research-Extensive Institution," Corbin M. Campbell and Alberto F. Cabrera actually tested a single institution with an intense case study to determine the construct and predictive validity for indigenous (non-transfer) graduating seniors at a mid-Atlantic research-extensive institution. Prompted by Porter's (2009) earlier work on the benchmarking of NSSE, where he began questioning the validity and reliability standards, two important measures, they found high inter-correlations among the benchmarks, low item loadings, and low reliability scores. Also, as a predictor of cumulative GPA for the case study, the authors determined that the NSSE benchmark model was not valid, at least regarding predictive validity. (In fairness, these assumptions are subject to the Porter caution about characterizing "high," "low," and in between.) Even Kuh (2009) has acknowledged: "Institution-specific analysis sometimes produce[s] factor structures different than the five benchmarks or clusters of effective educational practices that NSSE uses to report its findings" (p. 687).

Although one of the more promising features in the rise of rankings is the move to measure institutional effectiveness through student engagement, at the least, this move to the engagement regime should do no harm. The authors are particularly skeptical about the efficacy of this new approach and fear that one imperfect measure may be replaced by another, worse set of measures; they are particularly concerned about whether the five NSSE benchmarks will prove to be reliable and/or valid at the institutional level, especially as few studies have been undertaken on exactly these psychometrics, and those few studies found very weak measures. Campbell and Cabrera (2011) sum them up almost dismissively: The studies

found that (a) the construct validity of certain benchmarks was either marginal or poor, (b) the benchmarks did not appear to be strongly associated with important student outcomes, like GPA, and (c) the benchmarks were

highly intercorrelated: they appear not to measure distinct domains of student engagement. (p. 80)

If these findings remain true, it will be devastating for the use of these SSE measures, for many schools are using them, even with the clear limitations; and their use will become harder when the required major adjustments require schools, in essence, to surrender any over-time comparability as new adjustments are made. Worse, they could be proven to have been the cold fusion of institutional research, all hat and no cattle, and a tremendous investment never to pay off or to improve practice (Carle, Jaffe, Vaughn, & Elder, 2009; Gordon, Ludlum, & Hoey, 2008; Kuh et al., 2008; LaNasa, Olson, & Alleman., 2007; LaNasa, Cabrera, & Transgud, 2009; Pike, 2006).

Campbell and Cabrera are after big game:

> If the NSSE benchmarks are a valid measure of student engagement, they should be predictive of student learning across a variety of institutional types and student populations (i.e., have predictive validity). Additionally, a strong measure of institutional effectiveness must have construct validity, which in the case of NSSE, would be evidence that the five benchmarks measure five distinct, yet intercorrelated, domains of engagement. Evidence of predictive and construct validity gathered from single- and multi-institutional samples would support the claim that the NSSE benchmarks are a valid method for ascertaining institutional effectiveness. (pp. 80–81)

The five NSSE Benchmarks of Effective Educational Practice are:

- Level of Academic Challenge (LAC)
- Active and Collaborative Learning (ACL)
- Enriching Educational Experiences (EEE)
- Student-Faculty Interaction (SFI)
- Supportive Campus Environment (SCE)

The authors did not have to add that lack of such evidence could be devastating, as both institutional users and the scholars who evaluate and study the process could lose confidence in the measures. And all of us who care about this phenomenon would be back at square one in search of the key to all mythologies, the chimerical perfect measure, and the political silver bullet.

In a very exacting and careful fashion, these authors sift through the five benchmarks very thoroughly—and find them wanting. They find several of the important benchmarks to be weak markers, and they comb through literally dozens of in-house and more arm's-length reviews on the psychometrics of the NSSE, and discover especially poor reliability and validity features. While they may be questioned for having used the GPA as their student outcome measure, given how much variability there may be across majors, college subunits, and the like, it is surely a reasonable and widely

used criterion. Their findings seem well constructed and intuitive. Cabrera is surely among the most accomplished of the scholars in this field, with no discernable biases or axes to grind. The recommendations are very reasonably argued, and the authors are very judicious in their tone and approach:

> Our findings question the extent to which NSSE benchmarks are a universal tool for appraising institutional quality, and whether they predict such student outcomes as GPA. We echo Gordon, Ludlum, and Hoey's (2008) advice to institutional researchers and policymakers. They should carefully examine the extent to which the five NSSE benchmarks are reliable and valid for their own institutional contexts before committing themselves to major organizational changes. If each of the five benchmarks does not measure a distinct dimension of engagement and includes substantial error among its items, it is difficult to inform intervention strategies that will improve undergraduates' educational experiences. For example, if it is unclear what the EEE benchmark actually measures, the interventions could be targeting the wrong precursor of learning.
>
> Additionally, our results, in consonance with research at other institutions, may suggest the refinement of NSSE benchmarks to be more reliable and valid measures. NSSE plans to launch an updated version of the survey in 2013, dubbed, "NSSE 2.0" (NSSE, 2011). Our research would suggest that the NSSE researchers pay special attention to the psychometric properties of the revised benchmarks at the institutional level to ensure that NSSE 2.0 reaches its full potential as an instrument to measure engagement and student learning across institutions. (p. 97)

In addition to publishing traditional articles in the research literature, NSSE scholars have taken the step of maintaining a website "psychometric portfolio" to show how robust its measures of engagement and psychometrics are: "As part of NSSE's commitment to transparency as well as continuous improvement, we routinely assess the quality of our survey and resulting data, and we embrace our responsibility to share the results with the higher education community" (NSSE, 2010b). The website lists a great deal of information about the various SSE enterprises, as well as a growing research library with studies of reliability, validity, and other quality indicators, as one might expect from the largest and best-established college student assessment project in U.S. history. As I note in my conclusion, there are some subject matter gaps in its internal scholarship, and I will urge attention to these, several of which arise from this careful article.

The final article on this special theme is, unusually, by Amaury Nora, the editor of this journal, and his colleagues Gloria Crisp and Cissy Matthews: "A Reconceptualization of CCSSE's Benchmarks of Student Engagement." (Disclosure: All three were my doctoral students at the University of Houston, and I chaired Nora's dissertation committee and served on Crisp's when, 20 years later, she was our student). His appearance in this volume shows

both his devotion to this data base and getting it right, and his stature in the field. He and Cabrera (also formerly of the University of Houston) began engaging the work of Tinto, Bean, Pascarella, and Terenzini (1991) years ago, particularly noting the differences between White student constructs and those for students of color, particularly Latinos (Nora & Cabrera, 1996; Nora, Cabrera, Hagedorn, & Pascarella, 1996; Cabrera, Nora, Terenzini, Pascarella, & Hagedorn, 1999; Nora, 2001–2002).

Nora, Crisp, and Matthews (2011) have usefully reconceptualized the underlying model design of the CCSSE model, and this is an important contribution—at least as important as the more psychometric objections raised in the other articles. They proffer this suggestion:

> The proposed model theoretically incorporates the behavioral and psycho-social perspectives underlying the nature of student engagement and therefore represents a much more holistic view of the engagement of students on campus. The behavioral aspect represents the student's participation in educational activities both in and out of the classroom. This behavioral component of student engagement represents those actions or activities in which the student is involved. However, participation in a specific activity may be active or passive; it may involve sitting in a group that is using a collaborative learning approach but not mentally engaging in the activity with other students. This behavioral facet incorporates those activities included in CCSSE items and other behaviors such as joining in discussions, engaging in debates, questioning norms, and interacting with both faculty and students. (p. 126)

Under this new model of student engagement, they convincingly demonstrate that the unrestricted factor analysis of the data representing the CCSSE benchmarks represents a very different factor structure than the one claimed by the SSE architects. They also suggest that the hypothesized model would be more representative than the survey's benchmarks:

> In the reconceptualization model, the column labeled Academic and Social Integration represents the latent constructs associated with the behavioral and attitudinal aspects of the model. Both integration components guide the operationalization of observable variables—the items developed as indicators of behavioral and non-behavioral forms of academic and social integration or student engagement. Though not all-encompassing, the following examples provide a sense of items that could be developed to measure student engagement: finishing all assignments, meeting with a study group, having discussions with faculty, reading additional materials, feeling satisfied (measured by degree) with the instructional techniques utilized in the classroom, feeling validated in the classroom, perceiving that the institution and faculty are responsive to students, and developing an academic sense of belonging. Examples of the aspects of social integration or engagement include such variables as attending social activities, participating in student government or social enti-

ties on campus, developing friendships from study groups, becoming aware of diversity on campus, becoming part of a student support group, becoming a member of a social network, feeling that there is a fit between the institution and the student, developing an allegiance to the institution, and feeling there is a sense of tolerance on campus from faculty and other students. (pp. 126–27)

Conclusion and Suggestions to Professor Kuh

On the NSSE website, the validity section of the NSSE portfolio includes studies of seven types of validity (response process, content, construct, concurrent, predictive, known groups, and consequential) and three dimensions of reliability (internal consistency, temporal stability, and equivalence). The authors in this special issue fix their critical gaze even upon the discourse evident in this extensive website:

> It is clear that researchers at NSSE have conducted a wide array of psychometric studies on the properties of NSSE, but there are two glaring omissions. First, they have not reported construct validation of the five benchmarks of effective educational practices. Second, they cite no research examining how well the benchmarks hold true for individual institutions. (Campbell & Cabrera, p. 85)

Ouch!

The Review has invited Professor Kuh and his colleagues to respond to these papers in a forthcoming issue. George: Run, do not walk, to accept this solicitation. I certainly hope that you and your collaborators will carefully and fully respond to the many items raised by Cabrera, Porter, Dowd, Nora, and Crisp, and their graduate students. I would first suggest that any response grapple with a much more—shall I say—heterogeneous literature? At the exact same time I began this preface, in April 2011, I received my personal copy of *Research in Higher Education*, vol. 52, no. 3 (May 2011), in which Kuh and two frequent collaborators published "An Investigation of the Contingent Relationships between Learning Community Participation and Student Engagement" (Pike, Kuh, & McCormick, 2011, pp. 300–322) Interestingly, after finding ("eureka") that the results were "positively and significantly related to student engagement," the authors conceded:

> Analyses also revealed [that] there was substantial variability across institutions in the magnitude of the relationships between learning community participation and first-year students' level of engagement. Although institutional characteristics accounted for some of the variability across institutions, a substantial amount of the variability in engagement-learning community relationships remained unexplained. (p. 300)

Ten years after the NSSE first appeared, this candid acknowledgement is a major concession.

But even more interesting is how the authors situate their work in the overall literature of what they call "the current research" (p. 301). In their bibliography, the authors cite 75 works, certainly on the high side for this genre, but a careful count shows that Kuh authored 18 of these pieces, while Gary Pike authored 10 more. Even so, these figures undercount the totality of their work as reference materials, as the bibliography includes several SSE studies and case studies that had no identifiable authors. While it is a good thing that they had several Asian or Asian American collaborators, there is virtually no work by Latino or African American scholars. Marvin Titus is cited for one article; Nora is cited once, in a multi-authored book chapter, whose lead author is Ernest Pascarella. I might be undercounting, as the race of all of the authors is not perfectly discernable, but I recognize most of the persons cited, and I would just urge skeptical readers to contrast this provenance with the work cited by the authors in this single volume. Their references bristle with the parallel universe of minorities writing about a predominantly minority topic in a larger White world. While the entire point is that these scholars are deeply reading and discerning the work of Kuh and NSSE, it is not clear that the reverse is at work. Contrast these reading lists, and draw your own conclusions.

Second, I urge anyone from NSSE who does respond to these respectful critiques: Do not be thin-skinned or dismissive. I urge the same sensibility toward this collection of papers that their authors have shown to NSSE. All these critical authors give it its due, and the smartest thing that the Indiana enterprise could undertake is to draw these scholars into the tent, taking their suggestions seriously and in the spirit they are clearly intended. If they did not believe in the long-term efficacy and possibilities in NSSE and the various related projects, they would not be spending their considerable critical acumen on the topic.

Finally, while the CCSSE, NSSE, and LSSE have all had considerable market success, serious questions are being raised here. If the assessment movement is to become more than voodoo economics, especially with all the minority and first-generation students entering our colleges, the one size cannot fit all. These scholars have provided a detailed roadmap for improvement and have questioned all the topographical markers. They have also seriously questioned the instruments we use to read maps and, more crucially, how we draw maps. I look forward to this—well, dare I say it?—engagement.

REFERENCES

Astin, A. W. (1993). What matters in college? *Four critical years revisited.* San Francisco: Jossey-Bass.

Astin, A.W. (1999). Student involvement: A developmental theory for higher education. *Journal of College Student Development, 40,* 518–529.

Cabrera, A., Nora, A., Terenzini, P., Pascarella, E., & Hagedorn, L. (1999). Campus racial climate and the adjustment of students to college: A comparison between White students and African-American students. *The Journal of Higher Education, 70*(2), 134–160.

Campbell, C. M., & Cabrera, A. F. (2011). How sound is NSSE? Investigating the psychometric properties of NSSE at a public, research-extensive institution. *Review of Higher Education, 35*(1): 77–103.

Carle, A. C., Jaffe, D., Vaughn, N. W., & Eder, D. (2009). Psychometric properties of three new national survey of student engagement based engagement scales: An item response theory analysis. *Research in Higher Education, 50,* 775–794.

Chickering, A. W., & Reisser, L. (1993). *Education and identity.* San Francisco: Jossey-Bass

Dowd, A. C., Sawatzy, M., & Korn, R. (2011). Theoretical foundations and a research agenda to validate measures of intercultural effort. *Review of Higher Education, 35*(1): 17–44.

Goodwin, L. D., & Leech, N. L. (2003). The meaning of validity in the new standards for educational and psychological testing: Implications for measurement courses. *Measurement and Evaluation in Counseling and Development, 36*(3), 181–192.

Gordon, J., Ludlum, J., & Hoey, J. J. (2008). Validating the NSSE against student outcomes: Are they related? *Research in Higher Education, 49,* 19–39.

Gratz v. Bollinger, 539 U.S. 244 (2003).

Grutter v. Bollinger, 539 U.S. 306 (2003).

Guerra Thompson, S. (2006). Evading Miranda: How Seibert and Patane failed to "save" Miranda. *Valparaiso Law Review, 40,* 645–684.

Kuh, G. D. (2000a). *The NSSE 2000 report: National benchmarks for effective educational practice.* Bloomington: Indiana University Center for Postsecondary Research and Planning.

Kuh, G. D. (2000b). T*he national survey of student engagement: The college student report.* Bloomington: Indiana University Center for Postsecondary Research and Planning.

Kuh, G. D. (2009). What student affairs professionals need to know about student engagement. *Journal of College Student Development, 50,* 683–706.

Kuh, G., Cruce, T., Shoup, R., Kinzie, J., & Gonyea, R. (2008). Unmasking the effect of student engagement on first-year college grades and persistence. *The Journal of Higher Education, 79*(5), 540–563.

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

LaNasa, S. M., Cabrera, A. F., & Transgrud, H. (2009). The construct validity of student engagement: A confirmatory factor analysis approach. *Research in Higher Education, 50,* 315–332.

LaNasa, S. M., Olson, E., & Alleman, N. (2007). The impact of on-campus student growth on first-year student engagement and success. *Research in Higher Education, 48*(8), 941–966.

Museus, S. D., & Maramba, D. C. (2011). The impact of culture on Filipino American students' sense of belonging. *The Review of Higher Education, 34*(2), 231–258.

Nora, A. (2001–2002). The depiction of significant others in Tinto's "Rites of Passage": A reconceptualization of the influence of family and community in the persistence process. *Journal of College Student Retention: Research, Theory & Practice, 3*(1), 41–56.

Nora, A., & Cabrera, A. (1996). The role of perceptions of prejudice and discrimination on the adjustment of minority students to college. *Journal of Higher Education, 67*(2), 119–142.

Nora, A., Cabrera, A. F., Hagedorn, L., & Pascarella, E. T. (1996). Differential impacts of academic and social experiences on college-related behavioral outcomes across different ethnic and gender groups at four-year institutions. *Research in Higher Education, 37*(4), 427–452.

Nora, A., Crisp, G., & Matthews, C. (2011). A reconceptualization of CCSSE's benchmarks of student engagement. *Review of Higher Education, 35*(1): 105–130.

NSSE. National Survey of Student Engagement. (2008). *Origins.* Retrieved on April 11, 2011, from http://nsse.iub.edu/html/origins.cfm.

NSSE. National Survey of Student Engagement. (2010a). *NSSE timeline 1998–2009.* Retrieved on April 11, 2011, from http://nsse.iub.edu/pdf/NSSE_Timeline.pdf.

NSSE. National Survey of Student Engagement. (2011a). *Psychometric portfolio.* Retrieved from on April 11, 2011, http://nsse.iub.edu/_/?cid=154.

NSSE. National Survey of Student Engagement. (2011b). N*SSE 2.0 goes live in 2013.* NSSE E-News. Retrieved from on April 11, 2011, http://nsse.iub.edu/e-news/2011_February.cfm#anchor_3.

Nuñez, A. M. (2009). Latino students' transitions to college: A social and intercultural capital perspective. *Harvard Educational Review, 79*(1), 22–48.

Olivas, M. A. (1986). Financial aid and self-reports by disadvantaged students: The importance of being earnest. *Research in Higher Education, 25*(3), 245–252.

Pascarella, E. T., & Terenzini, P. T. (1991). *How college affects students: A third decade of research.* San Francisco: Jossey-Bass.

Pike, G. R. (2006). The convergent and discriminant validity of NSSE scalelet scores. *Journal of College Student Development, 47,* 550–563.

Pike, G. R., Kuh, G. D., & McCormick, A. M. (2011). An investigation of the contingent relationships between learning community participation and student engagement. *Research in Higher Education, 52*(3), 300–322.

Porter, S. R. (2009). *Do college student surveys have any validity?* Paper presented to the Annual Meeting of the Association for the Study of Higher Education. Vancouver, Canada. Reprinted *Review of Higher Education, 35*(1): 45–76.

Stinebrickner, R., & Stinebrickner, T. R. (2004). Time-use and college outcomes. *Journal of Econometrics*, 121, 243–269.

Tanaka, G. (2002). Higher education's self-reflexive turn: Toward an intercultural theory of student development. *Journal of Higher Education, 73*(2), 263–296.

Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition* (2nd ed.). Chicago: University of Chicago Press.

U.S. GAO. U.S. Government Accountability Office. (2010) Undercover testing finds colleges encouraged fraud and engaged in deceptive and questionable market-

ing practices. GAO-10-948T (August 2010). Retrieved on April 11, 2011, from http://www.gao.gov/new.items/d10948t.pdf.

U.S. Senate. (2010). Health, Education, Labor and Pensions Committee. *The return on the federal investment in for-profit education: Debt without a diploma.* September 30, 2010. Retrieved on April 11, 2011, from http://harkin.senate.gov/documents/pdf/4ca4972da5082.pdf.

Wolf-Wendel, L., Ward, K., & Kinzie, J. (2009). A tangled web of terms: The overlap and unique contribution of involvement, engagement, and integration to understanding college student success. *Journal of College Student Development, 50*(4), 407–428.