

DRAFT – ASHE 2009 – *Please do not cite or reproduce.*

Tracking the Impact of Assessment: Studying Evidence-Based Improvement in Colleges and Universities
Jillian Kinzie and Alexander C. McCormick
Indiana University Center for Postsecondary Research

The history of assessment in higher education spans more than two decades. The National Institute of Education's 1984 report, *Involvement in Learning*, included the following recommendation: "Student evaluations of academic programs and the learning environment should be conducted regularly. The results should be widely disseminated as a basis for strengthening the quality of undergraduate education" (p. 61). The report also emphasized the goal of improving educational quality and challenged institutions to adopt organizational strategies, policies, and processes that would maximize the use of evidence to inform institutional improvement. In the decades since this clarion call, commitment to and professionalization of assessment practice in higher education has steadily increased and more recently, there is greater press to demonstrate educational effectiveness and that assessment activities impact institutional improvement initiatives.

Although institutions of higher education have increased their assessment activities (Banta, Jones, & Black, 2009; Kuh & Ikenberry, 2009), there is little evidence of the implementation of the phase of the assessment loop that involves following up to see whether the changes undertaken have produced improvements (Banta, Jones, & Black, 2009; Suskie, 2004). Given the press to demonstrate quality in colleges and universities and maturation of the assessment movement, there is reason to turn greater attention to the extent to which institutions are closing the assessment loop.

This study grows out of concerns about improving undergraduate education, the extent to which assessment activities and the results of educational reform efforts can be identified and documented, and the need to understand the patterns and trends in changes in assessment results over time. Specifically, it is motivated by a unique opportunity to (a) identify cases of institutional improvement focused on effective practices in undergraduate education at a wide range of colleges and universities, (b) examine the changes to gain insights into the nature and patterns of change in undergraduate education, and (c)

ultimately, to inquire into how change was effected. Assessment of student learning and other measures of educational effectiveness has tremendous diagnostic value. However, the true test of an effective assessment program is if it leads to appreciable difference in educational quality. Moreover, college and university leaders want to know if the improvements they invest in are achieving desired results, and leaders of other institutions are hungry for models of successful change. This study informs important but more general questions about improvement in higher education by identifying if institutional assessment results change, and if so, what are the patterns of change and the circumstances that enabled intentional change to take root and thrive.

Review of the Literature

This review of the literature addresses the topics of change in institutions of higher education, the linkage between assessment and improvement, and the concept of measuring change over time.

Change and Reform in Higher Education

Change in institutions of higher education is as dynamic and complex as it is in any other organization (Kezar, 2001). Hearn (1996) asserted several propositions about transforming U.S. higher education, including insights that organizational politics and institutional culture are important considerations for change, and that transformation is resource-dependent. Many studies of change in higher education have demonstrated the existing structures of colleges and universities, including loosely coupled systems, shared governance, and employee commitment, affect the change process. In addition, change models for institutions of higher education include an array of variables such as the environment, political climate, the commitment of leaders, and the involvement of the whole system (Kezar, 2001). Implementing large-scale, transformational change in colleges and universities is difficult, due in part to their complexity and also to their governance practices (Cuban, 1999; Birnbaum, 2000). Change models in institutions of higher education are complex, take many forms, and are highly dependent on the type of change being initiated.

Improvement and change in institutions of higher education took on a more pointed perspective with the beginnings of the assessment movement. By definition, assessment is an ongoing process aimed

at understanding and improving student learning, that requires the gathering of evidence to evaluate performance, and the use of results to document and improve performance (Angelo, 1995). One of the founding principles identified in the groundbreaking *Involvement in Learning* report (Study Group on the Conditions of Excellence, 1984) highlighted the value of information and the press to improve the quality of undergraduate education. This report pointed out that in the course of transforming itself into a mass industry, America's colleges and universities had lost sight of the conditions for excellence in undergraduate education. It critiqued the notion that quality was about reputation and resources, and instead asserted a conception that was based in evidence of quality, and more importantly, in ensuring effective educational practices such as setting high expectations, providing involving settings and forms of pedagogy, and giving students continuous meaningful feedback.

Other notable national reports, including "Making Quality Count in Undergraduate Education" (1995), emphasized the need for improvement in undergraduate education and for evidence of quality. Generalizations about assessment formulated as "The Principles of Good Practice for Assessing Student Learning" promoted the importance of change in the statement that "assessment is most likely to lead to improvement when it is part of a larger set of conditions that promote change" (American Association for Higher Education, 1992, p. 3). Angelo (1996) concluded that assessment had replaced new instructional technologies and financial constraints as the prime driver for the promotion of transformation in colleges and universities.

The current quality improvement and reform agenda in higher education has grown from the roots of the quality emphasis of the early assessment movement. However, over the last decade, it has been impelled by more rigorous regional accreditation standards for evidence of student learning outcomes and more recently, a real concern about federal government intervention if institutions fail to provide convincing evidence of their effectiveness (Eaton, 2008; Kuh & Ikenberry, 2009). Recent initiatives including the Voluntary System of Accountability (VSA), developed through a partnership between the American Association of State Colleges and Universities (AASCU) and the Association of Public and Land-grant Universities (APLU) as a tool to help institutions demonstrate accountability, measure

educational practices and outcomes, and assemble information that is accessible, understandable, and comparable, was a direct response to the threat of governmental intervention in higher education. These pressures present a challenge to the assertion that real institutional change – change that derives from commitment and not just compliance – comes from within an institution, one that has created a culture characterized by shared power, imagination, transparency, a spirit of generativity, and mutual respect (Kezar & Eckel, 2002; Schein, 1999; Weick, 2005).

Linking Assessment and Improvement

Assessment and improvement are linked practices. Quite simply, successful assessment is directed toward improvements in teaching, student learning, academic and support programs and institutional effectiveness (Banta, Lund, Black, & Oblander, 1996). In the early 1990s, only about half of college and university administrators surveyed believed that assessment offered a way to improve undergraduate education (El Khawas, 1995). The perception of assessment to improve is more favorable now, in part due to the growing emphasis from regional accreditation groups on assessing student learning outcomes and showing evidence of improvement. In a 2009 survey of chief academic officers (Kuh & Ikenberry, 2009), the push from accreditors was identified as the chief driver for assessment, followed closely by a commitment to improvement. Wide spectrums of institutions, with exemplary practice in assessment, have long histories of using information to guide decision-making and to make improvements in undergraduate education (Banta, Jones & Black, 2009; Banta, Lund, Black, & Oblander, 1996). However, Banta, Jones & Black (2009) were surprised that despite significant advances, including the use of technology to organize and manage the assessment process and increased reliance on assessment findings to make key decisions aimed at enhancing student learning, the assessment for improvement activities at most of the profiled programs had only been underway for two to four years.

The final phase of any assessment and improvement plan involves using taking action on results and then going back to measure improvement, commonly referred to as “closing the assessment loop.” This is perhaps the most difficult phase of any campus assessment plan. Institutions of higher education find it difficult to convert data into action and even more challenging to measure the impact of

improvement initiatives. That said, most assessment activities are associated with some sort of campus improvement or change. Banta, Jones & Black (2009) found that institutions profiled for their good practice in planning, implementing, and sustaining assessment, reported that assessment was most related to improving *processes*, such as better teaching methods, enhanced advising approaches, better faculty development programs, or increasing resources for a program. However, they reported that only six percent of the cases that used assessment to improve were able to say that student learning had actually improved.

Measuring Change and Improvement

The lack of evidence for assessment resulting in actual improvement in student learning is consonant with statements by Bob Connor, President of the Teagle Foundation, who has been insistent about promoting the use of assessment to improve student learning (Connor, 2006). Connor has expressed genuine concern and frustration about the lack of real evidence of improvement in student learning. Another newly formed organization, New Leadership Alliance for Student Learning and Accountability (NLA/SLA), has as its primary goal to develop a mechanism for harnessing and directing collective, sustained, strategic action to improve student learning in American colleges and universities. The NLA was established based on nearly two years of discussion among heads of the major higher education associations, foundation representatives, and other education leaders to scale up efforts to support and improve assessment of, and accountability for, student learning in higher education. These organizations and others, including the Association for American Colleges and Universities (AAC&U) have extolled the importance of measuring educational improvement.

Calls to improve the quality of undergraduate education and increased attention to both the processes for assessing student learning and the results of those assessments have prompted many colleges and universities to launch significant educational reform initiatives (Eaton, 2001; 2008; Kezar & Eckel 2002). In many cases, assessment information—diagnostic evidence pointing to what is working and what can be improved—informs these reform initiatives. Information from assessment efforts can guide the improvement process in two ways, by focusing attention on desired transformations and by

providing an occasion to reflect on progress. Repeated assessments over time thus afford an opportunity to evaluate progress and feed these results back into improvement initiatives, thereby closing the assessment loop (Maki, 2004). Assessment appears to facilitate institutional transformation, yet few studies examine before-and-after assessment results to substantiate these transformations (Keup, Astin et al. 2001; Jeffery 2008). Also missing are qualitative case studies that provide comprehensive information about the process of institutional change (Astin, Lindholm et al. 2001; Keup, Astin et al. 2001).

Clearly, the issue that is central to advancing assessment and reform efforts in higher education is how to convert assessment results into real change that can be measured for its impact on student engagement and learning. The technology of gathering assessment data has expanded. However, more evidence is needed to show that quality is being sustained or that improvements have had the desired effect on educational quality.

Theoretical Framework

Assessment and improvement in higher education provides a practical orientation to our study. However, our work is more generally informed by treatments of organizational learning and intentional change in organizations (Dill, 1999; Fiol & Lyles, 1985; Garvin, 1993; Milam, 2005), with a particular focus on change in institutions of higher education. Organizational learning provides the broad frame for inquiry into how institutions acquire and interpret information, make choices, and then interpret their experience, while the literature on change provides a sharper focus on goal-directed change (e.g., performance improvement). The literature on organizational learning is wide-ranging, but generally holds that organizational learning occurs when “acquiring, distributing or interpreting information” leads to a change in the range of potential behaviors (Huber, 1991, p. 89).

From an evaluation perspective, organizational learning involves intentional, goal-directed change that uses information or feedback to make changes (Torres & Preskill, 2002). Argyris and Schön’s (1978) assert that it is “a surprising mismatch between expected and actual results” that stimulates change. Our assumption is that assessment data provide such disconfirmation, and thus motivate action to render actual experience more consonant with “expectations or hopes.” Finally, recent projects examining

institutional change in higher education, including the Diversity Scorecard (Bauman, 2005), also provides a frame for examining how assessment results inform change within colleges and universities, and also to understanding the processes that produced successful outcomes.

Research Design & Method

To advance our understanding of using assessment results to improve educational effectiveness and the extent to which institutions have closed the assessment loop, we looked to a 10 year old assessment project involving more than 1,400 institutions as a viable data source for this study. Each year, several hundred baccalaureate-granting institutions participate in the National Survey of Student Engagement (NSSE) to assess the extent to which their undergraduate students (specifically, first-year students and seniors) engage in effective educational practices. Designed as a tool that institutions can use to document and evaluate the quality of undergraduate education, NSSE surveys random samples of first-year and senior undergraduates to assess the time and effort they put into their studies and the extent to which they are exposed to and engage in empirically-proven educational practices (Chickering & Gamson, 1987; Kuh, 2001). NSSE also assesses how an institution organizes learning opportunities and services to promote student learning and success. Because NSSE focuses on actionable, diagnostic information about teaching and learning, the results can be used to identify strengths and shortcomings in the quality of undergraduate education, and this information, in turn, can inform plans for improvement.

The full range of baccalaureate degree-granting institutions have participated in NSSE, and as of 2009, more than 600 institutions have participated at least three times, with many having participated four, five, or more times. The availability of institutional results over several administrations affords a unique opportunity to examine change in educational practice over time, to identify institutions where positive changes resulted from intentional improvement efforts, and to investigate what contributed to the success of these efforts. In short, NSSE provides a window into evidence-based improvement and organizational learning.

From the outset, one of NSSE's principal goals has been to provide participating colleges and universities with diagnostic, actionable information that can be used to improve undergraduate education.

Steady growth in the number of participating institutions, large numbers that administer the survey on a periodic basis, and a decade of collecting evidence, combine to make 2009 an opportune moment to examine multi-year NSSE data for evidence of change at the institution level. In addition, NSSE has been cataloguing stories about using results to improve undergraduate education since its inception. Accounts from more than 400 baccalaureate granting institutions detail a range of uses. For example, using results to gain insight into what contributes to student persistence, monitoring students' participation in good educational practices, sharing unsatisfactory results about the quality of co-curricular experiences with Board of Trustees to stimulate greater investment of resources in campus life, reformulating general education, and studying students' participation in diversity experiences, just to name a few. Many of these accounts speak to intentional efforts to use assessment results to promote change. However, most of these stories discuss broad improvement goals, or changes in institutional practice. Far fewer institutions track evidence of specific changes in student engagement and most institutions are only at the early stages of considering student engagement trends over time.

From its inception, NSSE has emphasized the diagnostic value of student engagement data and reports. Random sampling ensures the comparability of results among institutions, and reports to institutions show how students' in- and out-of class activities and experiences, as well as their perceptions of relationships and institutional emphases, compare with those of students attending other institutions in up to three customizable comparison groups. While benchmarking performance against peer institutions is the most common way that colleges and universities evaluate their performance, it is not the only way to do so. Another informative way to understand performance is to monitor change or stability in an institution's results over time. How does current performance compare with that of two, three, four, or more years ago? What is the trend? How do these results comport with strategic priorities and improvement efforts? Institutions can use a single year's results to compare their educational effectiveness with that of peer institutions, but they can also use results from multiple administrations to benchmark against *themselves* over time. In this way, they can monitor progress toward their goals for undergraduate education and gauge the impact of improvement initiatives.

Taking advantage of the opportunity to review multiple years of NSSE results, we were able to explore the following research questions in this paper:

- Is it possible to detect evidence of change in multi-year results?
- What do the data tell us about whether campuses are realizing gains in student engagement?
- Are there identifiable patterns of improvement in multiple measures?

Additional questions about the nature and scope of change at the institution will be explored in a subsequent follow-up study involving case studies. For example, at institutions where the empirical trends in assessment data are confirmed by institutional informants to reflect organized improvement efforts, what accounts for the success of these efforts? What role did assessment results play in improvement efforts?

Approaches to Detecting Change

The first step for conducting our analysis of change was to examine the extent to which change can be detected in results. There are various ways to assess and interpret changes in student engagement scores over time, but we concur with Singer and Willet's (2003) suggestion on necessary features for analyzing change: three or more waves of data, an outcome measure that changes systematically over time, and a sensible metric for clocking time. Specifically, we tested different statistical methods to identify institutions where NSSE results show a significant change. We conducted this phase of the study in spring 2009 (Kinzie, McCormick, Korkmaz, & Buckley, 2009) and determined that it is possible to detect change in NSSE scores using t-tests, effect sizes, regressions. We then considered if change could be detected in multi-institutional results using Hierarchical Linear Modeling (HLM) due to the multiple institution context. We examined each measure for first-year and senior data separately, and allowed the random slope and random intercept to vary between institutions. We concluded that it is effective to use a two level HLM to produce adjusted school mean scale scores for each year, then in Excel, calculate R^2 values for each institution's scales. The criteria for identifying change is to have at least one R^2 value greater than .70 and medium effect size differences between the highest and lowest values.

As for the sensible measure, we focused on NSSE's benchmarks for effective educational practice and high-impact practices as measures of distinct educational practices. Our analyses include both student-level and institution-level information, but the unit of analysis in this study is the institution. To be clear, the “change” in this study is at the institution level, not changes (or growth) in students from first-year to senior year. Although four observations, each an estimate with uncertainty, is only a brief time series, we think it is a critical to examine if change is detectable. As indicated, we employed two different quantitative analyses, t-tests and effect-size, and hierarchical linear modeling (HLM), which examines institutions' slopes for each scale, to detect change. Using two approaches offers some corroboration in the analyses and exploring the use of simple t-tests provides an opportunity to explore the viability of institutional analyses of change.

Participants

For this analysis, we identified a group of U.S. institutions that participated in at least four NSSE administrations beginning in 2004 or 2005 and concluding in 2009 (years with stable survey items). The resulting group of institutions thus had from four to six distinct observations for first-year and senior students over the six-year time span. Data quality considerations (response rate, number of respondents, and sampling error) led us to exclude a small number of institutions so we could have confidence in each year's survey results. The analyses for first-year students are based on a group of 211 institutions, and senior analyses are based on 222 institutions. About three-quarters of these institutions had at least five data points. The group of institutions analyzed represented the diversity of U.S. higher education with respect to control, size, and Basic Carnegie Classification (Table 1).

Variables

Examining each institution's data across the multiple NSSE administrations, we looked for institutions with evidence of systematic change on five indicators (see Appendix A):

NSSE Benchmarks: (1) Level of Academic Challenge; (2) Active & Collaborative Learning; (3) Student-Faculty Interaction; (4) Supportive Campus Environment

High-Impact Practices: (5) Proportion of first-year students who participated in a learning community or in service-learning as part of a regular course, or proportion of seniors who participated in (a) a practicum, co-op, internship, or field experience, (b) research with a faculty member, (c) study abroad, (d) a culminating senior experience, or (e) service-learning.

This analysis relies on strictly comparable measures for each year examined, and uses the same measures for all institutions.

Because NSSE results are institutional estimates based on a sample of students, identifying change involves more than simply comparing average scores across administrations. To identify meaningful change, we asked three questions: (1) Is the difference between the first and last data points large enough that it is not likely to be due to chance variation between samples (that is, is it statistically significant)? (2) If significant, is it meaningful—is it large enough to be noticeable to an informed observer (in technical terms, does it achieve an effect size of at least .3)? (3) Does the pattern of four, five, or six data points provide a reasonable fit to a linear or curvilinear trend (that is, is the pattern of change reasonably systematic)? These criteria determined that aggregate institutions discussed in the results section. In addition, we were interested in examining more particular instances of institutional change and selected about 40 institutions with positive trend multi-year scores to examine as more in-depth as case studies. We reviewed these scores for trends and patterns in order to make some tentative propositions about institutional change.

Results

Consonant with our earlier exploratory study using three or four data points through 2008, we found that change can reasonably be detected in NSSE benchmark scores over time. Trend summaries indicate that of the institutions studied, 87 (41%) demonstrated a pattern of improvement in at least one of the criterion measures for first-year students, and 63 (28%) did so for seniors (Table 2). A number of institutions—13 percent of each group—exceeded our modest criterion for meaningful change by a wide margin (that is, an effect size of at least .5). The percentage of institutions whose benchmark and high-

impact practice scores *declined* across multiple administrations was trivial (five institutions for one measure and from zero to two on each of the remaining ones). Of course, this means that the majority of institution results showed stability in scores over time. With regard to our interest in detecting change, these findings suggest that change is possible.

As Table 2 illustrates, comparing public to private institutions showed some difference within first-year scores, with public institutions having a higher proportion of improved scores. Using broad categorizations of undergraduate enrollment, results show that first-year student engagement scores at institutions with small undergraduate enrollments (fewer than 2,500) captured the fewest institutions (36%) compared to large enrollment institutions (5,000-9,999), in which 46% of the institution scores changed over time. Among the institutions with changes in senior scores, the highest proportion (38%) was found among very large enrollment (10,000 or more) institutions. The smallest percentage (24%) was within the small enrollment institutions. We aggregated institutions into Basic Carnegie Classifications and found that change was detected across all types. However, the Doctorate-granting institution category had the highest percentages (52% of first-year scores, and 34% of seniors) of institutions with any improvement trend for both first-year and senior scores.

Some measures had evidence of improvement at more institutions than did other scales. Among the first-year student benchmark scores, the scales with the most institutions were active and collaborative learning and student-faculty interaction. The fewest institutions were found in the supportive campus environment benchmark. Overall, the proportion of institutions that showed change in senior benchmark scores was pretty evenly distributed across the scales, with the greatest proportion in supportive campus environment (see Figure 3) and high-impact practices. These results suggest that the benchmarks that showed the highest proportion of change in first-year student engagement scores may be more amenable to improvement, or alternatively, that more institutions have targeted the first-year experience for improvement. Likewise, trends in senior student benchmark scores suggest that supportive campus environment and high-impact practices may also be most amenable to improvement, and that high-impact practices may be a focused area for improvement in the senior year. An examination of benchmark

combinations – institutions showing increases in at least two benchmarks -- indicates that the active and collaborative learning and student-faculty interaction among first year scores was the most popular pairing, and that student-faculty interaction and supportive campus environment was the most frequent combination for senior scores. The incidences of first-year and senior score improvements were rare with only the active and collaborative learning benchmark showing a combined impact.

Figures 1-3 display examples from three benchmarks of positive trends at selected institutions with varying levels of initial performance. These graphs depict the nature of improvement at individual institutions. To explore institutional improvement issues further, we next examined multi-year results at approximately 40 institutions that showed positive trends. This case-level examination affords a deeper level of analysis. By examining scores across each of the institutions with positive trends we were able to begin to construct stories about change in scores and the relationship between all benchmarks from an institutional point of view. The case study portion of this study is not yet complete so only a few examples from the institutional accounts will be featured in this paper.

In sum, although the magnitude and pattern of change varied, we found evidence of improvement on at least one measure at a great number of institutions. In addition, we found some intriguing patterns of change that are suggestive of systematic improvement efforts. For example, some institutions showed improvement on the same benchmark and scale such as active learning, and high-impact practices, for both first-year and senior students, while others showed improvement across multiple scales for just one of the two populations. We also found that some measures had evidence of improvement at more institutions than did other scales, suggesting either that some may be more amenable to intervention or that some tap areas of greater institutional interest or attention. Although five years may be a short time in which to expect change to take root in order to show in institution-wide results, this study demonstrates that NSSE data provide potentially fertile ground for examining the relationship between assessment and intentional improvement in undergraduate education.

Discussion

Many colleges and universities are engaged in meaningful assessment activities and their results inform institutional improvement activities. However, as Banta, Jones and Black (2009) conclude, too few have examined whether changes undertaken have produced desirable improvements. This study provides some evidence of the potential for institutions participating in NSSE, one straightforward assessment tool, to use their multi-year assessment results to detect evidence of change in student engagement results. Indeed, measuring the effect of institutional change is difficult. However, our quantitative examination using two different methods to identify change over time, demonstrates that gauging change is possible. For each selected NSSE scale (e.g., NSSE benchmarks, and other smaller scales including, active learning, diversity experiences, interaction with faculty outside of class, high impact practices), at least 30 or more institutions were identified having a positive change in their scores. Both simple methods using t-tests and effect sizes and more complex HLM analyses seem to be effective approaches to detecting change. This study offers a foundation for advocating different approaches to measuring change using multi-year analysis of NSSE results and offers some confirmation of the capacity to detect evidence of campus change initiatives in NSSE results.

The potential for examining change over time at the institution-level is an important consideration for campus assessment activities. The ability to detect with confidence that a change in a specific dimension of student engagement occurred is a vital piece of information for leaders of campus assessment and improvement efforts. Moreover, although it can be challenging to isolate the impact of a particular improvement, the examination of results over time affords an institution one gauge of the effectiveness of their reform initiatives. In this case, NSSE results could best be used as an indicator for identifying areas for more targeted follow-up and data collection to more fully estimate the impact of an initiative.

Although this study is focused on change, it is worth noting that examinations of multi-year results can also reveal stability and demonstrate consistency in educational quality. For many campus assessment plans it can be equally important to know that key indicators remain stable. Indeed, the

measures at nearly 59% of the institutions in this study were unvarying across four data points. In fact, several campus representatives at these institutions were pleased to learn that their results had not changed. One campus contact stated that she was relieved that despite a mandatory increase in class sizes, and some shifting in their undergraduate curriculum that could have upset the institutions good results, their scores in the academic challenge benchmark, particularly in writing items, remained strong, suggesting that their distinctive practices had not waned. For this institution, knowing that the modifications on campus had not cut into some of their most valued educational practices was important information.

Although the primary implication for these findings is aimed at institutional assessment and improvement activities, changes in student engagement scores over time also provide insights into the nature of the assessment cycle, particularly evidence of the link between assessment and improvement, patterns of change in educational practice, and greater understanding of the degree to which undergraduate education and more specifically, the quality of student learning is improving. The sections that follow examine the major themes of patterns of change among institutions and also across measures of effective educational practice and concludes with a discussion about challenges to conventional beliefs about institutional change in higher education.

Patterns of Change Across Measures. Institutions that sign on to administer NSSE usually do so with the intent to use results to demonstrate and improve the quality of undergraduate education. Multiple administrations provide the opportunity to measure progress and monitor improvement over time. Although institutional change can be difficult, our examination of change statistics across NSSE benchmarks and measures of high-impact practices demonstrate that it is possible to increase student engagement in effective educational practices.

The patterns of change across scales and between first-year and senior scores suggest several propositions about institutional change. First, *some benchmarks at specific points in the undergraduate program may be more amenable to change than others.* For example, one out of five institutions in the first-year analysis showed positive trends on two or more measures, and the most common combination

(found at 24 institutions, or 11% of the sample) was for the active and collaborative and student-faculty interaction benchmarks. The high proportion of institutions with improved scores in these two measures are likely due to the development of a robust body of knowledge about the policies, strategies, programs, and services that help first-year students make a successful transition to college and institutions creating an array of programs to challenge and support students in the first year. For seniors, the largest number of institutions showing upward trends was on the supportive campus environment benchmark and the proportion reporting involvement in high-impact practices. This might reflect department-level efforts to improve advising and academic support, or to provide opportunities for students to develop meaningful relationships with faculty and to engage deeply with what they are learning.

Second, *institutions showing same scale improvement for first-year students and seniors may reflect an institutional focus on this activity.* A parallel improvement agenda might be reflected in, for instance, increases in supportive campus environment scores for both first-year students and seniors. For at least one institution with this trend, the explanation for the increase was related to expanded, integrated, and better coordinated academic support services, and the creation of a “One Stop Shop” for student support services including financial aid, admission, bursar, registrar, and career services.

Third, *institutions showing a significant improvement across two scales within either the first-year or senior scores suggest class year focused improvement agendas.* This proposition is related to the first about the amenability of change, but is more reflective of an intentional improvement agenda. For example, the first-year trend results about active and collaborative learning and student-faculty interaction might reflect concerted efforts to enhance the first-year experience through projects such as the National Resource Center for the First-Year Experience and the Foundations of Excellence® project. Indeed, this combination of first-year student NSSE scores illustrate the impact of attention on the institutional structures – policies, practices, and procedures in the first year experience – that encourage new students to spend their time in educationally productive activities and return for the second year. The combination of improvement across two benchmarks is reminiscent of the finding that engaging educational practices are “sticky” (Kuh, Kinzie, Schuh, Whitt & Associates, 2005), in that they clump together. Once students

engage in meaningful ways with something that excites them, windows to other educationally purposeful activities open up. Thus, the impact on student success is multiplied when institutions create conditions and expose students to a variety of synergistic, effective, educational practices.

Although the examination of general trends in higher education is interesting, the more important considerations for change are at the institutional level. By examining each positive trend institution as a case study, we can get a more in-depth account of improvement over time, including some tentative explanations about institutional change and the potential for gauging the impact of reform. By plotting institutions' scores together for each benchmark scale we discovered change patterns. For example, some institutions seem to have a focused story about assessment data and change, as evidenced by their improved scores in a distinct educational practice. This is perhaps best illustrated by institutions with positive trends in active learning and high-impact practice scale scores for both first-year students and seniors over time. Accounts of using NSSE results at two institutions with this trend suggests that these institutions used their early NSSE assessment results (and likely other assessment information) on students' engagement in experiential learning activities to initiate curricular and co-curricular programs, services and enhanced pedagogical practice to address this dimension of the undergraduate experience. These institutions were also intentional about using follow-up NSSE administrations to gauge the impact of their initiatives. Institutional accounts of change can yield important insights into the change process.

Finally, college and university leaders and faculty members want to know if the improvements they invest in are achieving desired results. The examination of NSSE results over time provides one indicator of the extent to which reform efforts have had an impact on student engagement. In addition, several of these patterns of change are suggestive of systematic improvement efforts in higher education.

Challenging Beliefs about Undergraduate Education and Change. Our findings contradict some of the conventional wisdom about change in higher education. First, *the potential for improving performance is not limited to small institutions, private ones, or those with a low baseline level of performance.* We found patterns of systematic change at both public and private institutions, in every size

category, and in every Carnegie type (Table 2). Figures 1 - 3 display examples of positive trends at selected institutions with varying levels of initial performance.

Second, *increasing first-year student engagement in active and collaborative learning is possible at large public institutions*. Large classes, ranging in size from a run of the mill of 60 to upwards of 1,000 students, are common practice at large public institutions. The prevailing belief about teaching large classes is: anything you can do in a large class you can do better in a small one. However, plenty of research on effective pedagogy in large classes shows that there are ways to make large classes as effective as their smaller counterparts, such as, by getting students actively involved, helping them develop a sense of community, and providing frequent opportunities for in- and out-of-class collaborative learning (Johnson, Johnson & Smith, 1998; Michaelson, Knight, & Fink, 2002). As Figure 1 illustrates, although the large public university scored lower than the other institutions, its performance improved steadily and significantly.

Third, *contrary to beliefs that urban or commuter institutions can do little to increase student-faculty interaction, we found that 41% of the institutions that show positive trends on this benchmark for first-year students are urban institutions*. High quality, frequent interaction between students and faculty are activities typically associated with the student experience at small, private liberal arts colleges. This educational activity happens less often for students at urban or commuter institutions. However, our results showing that some urban and commuter institutions have improved scores on this dimension.

Fourth, *in contrast to the assumption that small, private institutions are most likely to make high-impact practices such as study abroad, undergraduate research and service-learning available to a significant proportion of their students, our results show that two thirds of institutions with improved results on student engagement in high impact practices are medium and large universities*. Given that all students, whatever their ethnic or economic background are more likely to persist in college and be engaged in high levels if they participate in high-impact practices (Kuh, 2008), it is an important dimension for institutional growth. However, despite the increases that these results show in student engagement in high-impact practices at a range of institutional types, it is concerning that a good many

students who've been historically underserved students in higher education are also underrepresented in these activities (Kuh, 2008). There is plenty of room to bring these programs up to scale at all institutions and thereby greatly increase their transforming effects on students' lives and learning.

Next Steps: Insights into Institution Specific Change. The results from this quantitative analysis of change suggest candidates for more in-depth, case study examinations of change activities. As we indicated, we are primarily interested in learning more about the process of change and how results were used to inform improvement from institutions with positive trend results. Although we have only just begun to assemble case study information, we expect to interview institutional contacts about their change process, what motivated their efforts, and the extent to which change has taken root, as the next step in this study. A few cases are already pointing to the importance of intentional action and the investment of resources in the change initiative. For example, a private liberal arts college in Georgia with positive trend scores in the first-year active learning scale was focused on improving this dimension of the undergraduate experience. Specifically, they invested in increasing active learning experiences in a reinvigorated First Year Seminar course with an emphasis on integrated writing across first year courses culminating in the creation of a first year writing portfolio that is shared with classmates and faculty in the program. The institution had been using NSSE results and other assessment data to regularly monitor the impact of their initiatives on both educational processes and student learning outcomes. Another case study effectively illustrates the importance of leadership and a systematic approach to using assessment results to guide improvement. A private, urban institution launched a study of the quality of service delivery and quality of campus relationships. Administrative leaders were already attuned to the need to reduce bureaucracy and develop more "one-stop services" and an analyses of multiple years of NSSE results on the quality of campus relationships scale plus some satisfaction data, reinforced this need. Student services received additional attention, particularly the registrar, bursar and financial aid operations, and were later restructured and renamed the office of student assistance. Follow up assessment results revealed positive changes scores.

Additional case study findings are suggesting more insights into the value of assessment results and of specific interventions to promote systematic use of data, the identification of internal and external influences in the processes associated with change, and notions about organizational structures and leadership factors that promote or obstruct these efforts. This work will illustrate factors that contributed to positive outcomes, explain the background of these situations, what happened and why, and explain why some innovations were particularly effective.

One of the practical implications for the hundreds of baccalaureate granting institutions that have participated in multiple NSSE administrations is to think more about exploring their results over time. NSSE encourages institutions with results from multiple NSSE administrations to examine patterns of change in their student engagement results. Some may want to look for changes in the nature of student engagement, some will investigate possible trends, while others will be keen to evaluate the impact of specific campus initiatives. Finally, while gross measures like the NSSE benchmarks, provide a reasonably good measure to track changes in student engagement over time, most institutions will want to identify more precise measures specifically associated to institutional improvement initiatives. For example, a specific improvement initiative to promote peer-to-peer learning in the first year experience deserves more precise measures than the broad indicators in the active and collaborative learning benchmark. Institutions interested in assessing change on campus are advised to look beyond their benchmark scores and consider changes in NSSE items or other scales, or to merge their NSSE results with other student information and outcome measures for more fine tuned or more complex examinations of change in student engagement over time.

Although this study is an important step in tracking the impact of assessment findings, it is clear that NSSE is just one source of assessment results. Results from other assessments may better detect different evidence or be more sensitive measures of campus change initiatives. In addition, while this study employs a retrospective design, the advantage of this approach is that it guarantees the opportunity to study positive change. Lessons from the case study portion of this study have the potential to make important contributions to our understanding of how educational organizations learn from assessment

efforts, use those results to inform reform initiatives, and ensure the success of those initiatives. The following questions will be explored in the next phase of this research: What accounts for the observed positive changes in student engagement at the subject institutions? What specific activities were undertaken to build consensus about the nature of the problem(s), the need for positive change, and the appropriate ways to effect change? What role did assessment data play in change initiatives? What specific institutional practices, policies, documents, and actions were invoked in the use of assessment data in campus change initiatives? To what extent were subsequent assessment data used to assess progress? Who were the key players in the change process? What activated their participation, how did specific actors or coalitions facilitate or resist change initiatives, how was resistance overcome?

Higher education institutions exist today in a tough climate of high expectations, suspicion about value, and diminished resources. Greater conversion of assessment results into measurable improvements in student learning, and more evidence showing that quality is being sustained or that improvements have had the desired effect on educational quality, are necessary to advance assessment and reform efforts in higher education. There is perhaps no better time than now to tell an evidence-based story about institutional improvement and for colleges and universities to effectively demonstrate their impact on student engagement and learning.

References

- American Association of Higher Education. (1992). *Principles of good practice for assessing student learning*. Washington, DC: American Association for Higher Education.
- Angelo, T. A. (April, 1996). Transforming Assessment: High Standards for Higher Education. *AAHE Bulletin*, 48 (8), 3- 4.
- Angelo, T.A. (November, 1995). Reassessing (and Defining) Assessment. *AAHE Bulletin*, 48 (2), 7-9.
- Argyris, C., & Schön, D.A. (1978). *Organization learning: A theory of action perspective*. Reading, MA: Addison-Wesley.
- Astin, A. W., Lindholm, J. A., Walker, A. A., & Keup, J. R. (2001). Facilitating transformative change efforts through the use of assessment. In A. W. Astin & H. S. Astin (Eds.), *Transforming institutions: Context and process*. Los Angeles: Higher Education Research Institute, UCLA.
- Banta, T. W., Jones, E.A., & Black, K.E. (2009). *Designing effective assessment: Principles and profiles of good practice*. San Francisco: Jossey-Bass.
- Banta, T.W., Lund, J.P., Black, K.E., & Oblander, F.W. (1996). *Assessment in practice*. San Francisco: Jossey-Bass.
- Bauman, G. L. (2005). Promoting organizational learning in higher education to achieve equity in educational outcomes. *New Directions in Higher Education* 131, 23-35.
- Birnbaum, R. (2000). *Management fads in higher education: Where they come from, what they do, why they fail*. San Francisco: Jossey-Bass.
- Chickering, A. W., & Gamson, Z. F. (1987). Seven principles for good practice in undergraduate education. *AAHE Bulletin*, 39(7), 3-7.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Commission on the Future of Higher Education. (2006). *A Test of Leadership: Charting the Future of U.S. Higher Education*. Washington, DC: U.S. Department of Education.
- Connor, W. R. (2006, March 31). From foxes to hedgehogs. *Inside Higher Ed*. Retrieved from <http://www.insidehighered.com/views/2006/03/31/connor>.
- Cuban, L. (1999). *How scholars trumped teachers: Change without reform in university curriculum, teaching, and research, 1890-1990*. New York: Teachers College, Columbia University.
- Dill, D. (1999). Academic accountability and university adaptation: The architecture of an academic learning organization. *Higher Education*, 3, 127-154.

- Eaton, J. S. (2008, March 24). The Future of Accreditation? *Inside Higher Ed*. Retrieved from <http://www.insidehighered.com/views/2008/03/24/eaton>
- Eaton, J. S. (2001). Regional accreditation reform: who is served? *Change*, 33(2), 38-45.
- Education Commission of the States. (1995). *Making quality count in undergraduate education*. Denver, CO: Education Commission of the States.
- Fiol, C.M., & Lyles, M. A. (1985). *Organizational Learning*. *The Academy of Management Review* 10(4), 803-813.
- Garvin, D.A. (1993, July-August). Building a learning organization. *Harvard Business Review*, 78-91.
- Hearn, J.C. (1996). Transforming U.S. higher education: An organizational perspective. *Innovative Higher Education*, 21(2), 141-154.
- Huber, G.P. (1991). Organization learning: The contributing processes and the literatures. *Organization Science* 2(1), 88-115.
- El Khawas, E. (1995). *Campus trends 1995: New directions for academic programs*. (Higher Education Panel Report No. 85). Washington DC: American Council on Education.
- Indiana University Center for Postsecondary Research. (n.d.) *Contextualizing NSSE effect sizes: Empirical analysis and interpretation of benchmark comparisons*. Retrieved from http://nsse.iub.edu/pdf/effect_size_guide.pdf.
- Jeffery, J. (2008). Institutions of higher learning and learning organizations: Annotated bibliography. Retrieved March 27, 2009, from http://www.teaglefoundation.org/learning/PDF/2008_LearningAB.pdf
- Johnson, D.W., Johnson, R.T., Smith, K.A. (1998). Cooperative Learning Returns to College: What Evidence Is There that it Works? *Change*, 20 (4), 26-35.
- Keup, J. R., Astin, H. S., Lindholm, J. A., & Walker, A. A. (2001). Organizational culture and institutional transformation. In A. W. Astin & H. S. Astin (Eds.), *Transforming institutions: Context and process*. Los Angeles: Higher Education Research Institute, UCLA.
- Kezar, A., (Ed.). (2001). Understanding and facilitating organizational change in the 21st century: Recent research and conceptualizations. [Special Issue]. *ASHE-ERIC Higher Education Report*, 28(4).
- Kezar, A., & Eckel, P. D. (2002). The effect of institutional culture on change strategies in higher education: Universal principles or culturally responsive concepts? *Journal of Higher Education*, 73(4), 435-460.
- Kinzie, J., McCormick, A., Korkmaz, A., & Buckley, J. (2009). Measuring Change: Using Multi-Year Analysis of National Survey of Student Engagement Results to Assess Educational Improvement. Presented at the Annual Forum of the Association of Institutional Research, Atlanta, GA.

- Kuh, G. D. (2001). Assessing What Really Matters to Student Learning: Inside the National Survey of Student Engagement. *Change*, 33(3), 10-17.
- Kuh, G.D., & Ikenberry, S. (2009). *More Than You Think, Less Than We Need: Learning Outcomes Assessment in American Higher Education*. National Institute for Learning Outcomes Assessment. University of Illinois.
- Kuh, G.D., Kinzie, J., Schuh, J.H., & Whitt, E.J. (2005). *Student success in college: Creating conditions that matter*. San Francisco: Jossey-Bass.
- Maki, P. (2004). *Assessing for learning: building a sustainable commitment across the institution*. Sterling, VA: Stylus Publishing.
- Michaelson, L. K. , Knight, A. B. & Fink, L. D. (eds.). (2002). *Team-Based Learning: A Transformative Use of Small Groups*. Westport, Conn.: Praeger.
- Milam, J. (2005). Organizational learning through knowledge workers and infomediaries. *New Directions for Higher Education* 131, 61-73.
- National Institute of Education. (1984). *Involvement in learning: Realizing the potential of American higher education*. Washington, D.C.: U.S. Department of Education.
- Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage Publications.
- Schein, E.H. (1999). *Process consultation revisited. Building the helping relationship*. Reading, MA: Addison-Wesley Inc.
- Schein, E.H. (1996). Kurt Lewin's change theory in the field and in the classroom: Notes towards a model of managed learning. *Systemic Practice and Action Research* 9(1), 27-47.
- Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.
- Suskie, L. (2004). *Assessing student learning: A common sense guide*. Bolton, MA: Anker Publishing Co.
- The Study Group on the Conditions of Excellence in American Higher Education. (1984). *Involvement in learning: Realizing the potential of American higher education*. Washington, DC: U.S. Department of Education.
- Torres, R.T., & Preskill, H. (2001). Evaluation and organizational learning: Past, Present, and Future. *American Journal of Evaluation*, 22(3), 387-395.
- Weick, K.E. (2005). Organization and failures of imagination. *International Public Management Journal*, 8(3), 425-438.

Table 1. Characteristics of the Multi-Year Institutions Analyzed in the Present Study

	First-year		Senior	
	Number	Percent	Number	Percent
Percent public	83	39%	90	41%
Size				
Small (fewer than 2,500)	86	41%	93	42%
Medium (2,500-4,999)	51	24%	53	24%
Large (5,000-9,999)	35	17%	36	16%
Very large (10,000 or more)	39	18%	40	18%
Basic Carnegie Classification (aggregated)				
Doctorate-granting	44	21%	44	20%
Master's Colleges and Universities	90	43%	95	43%
Baccalaureate Colleges	70	33%	76	34%
All others or unclassified	7	3%	7	3%
NSSE administrations between 2004 and 2009				
Four	54	26%	59	27%
Five	66	31%	68	31%
Six	91	43%	95	43%

NOTES: Cells contain column percents. The number of institutions varies by student population due to criteria for inclusion (i.e., differences in response rate or sample size between first-year and senior respondents).

Table 2. Institutions with Any Improvement Trend, by Selected Institutional Characteristics

	First-year		Senior	
	Number	Percent	Number	Percent
Total	87	41%	63	28%
Control				
Public	40	48%	23	26%
Private	47	37%	40	30%
Size				
Small (fewer than 2,500)	31	36%	22	24%
Medium (2,500-4,999)	23	45%	17	32%
Large (5,000-9,999)	16	46%	9	25%
Very large (10,000 or more)	17	44%	15	38%
Basic Carnegie Classification (aggregated)				
Doctorate-granting	23	52%	15	34%
Master's Colleges and Universities	32	36%	29	31%
Baccalaureate Colleges	29	41%	17	22%
All others or unclassified	3	43%	2	29%

NOTE: Cells contain the number and percentage of institutions with the indicated attribute that showed a pattern of improvement on at least one criterion measure.

Figure 1.

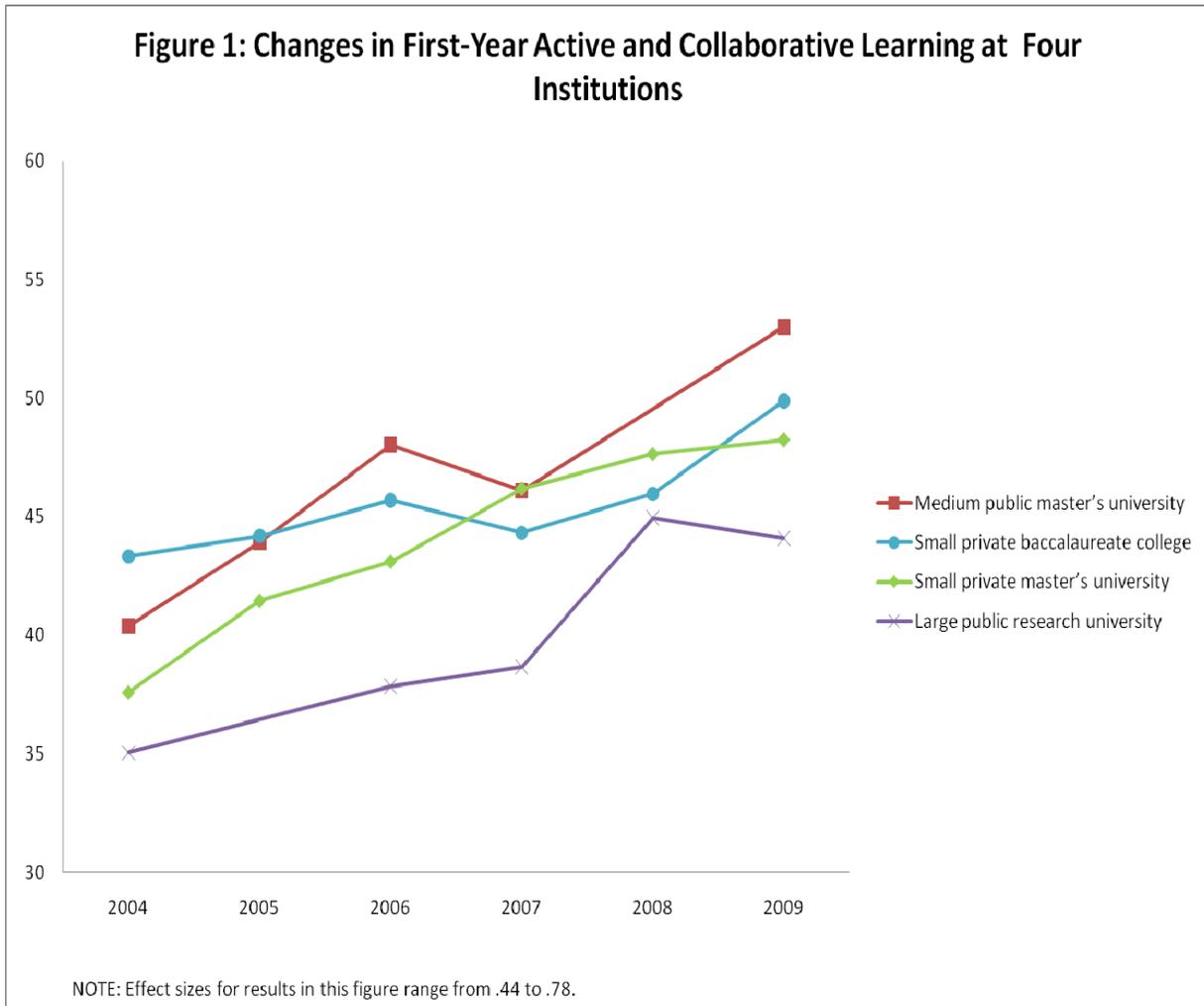


Figure 2.

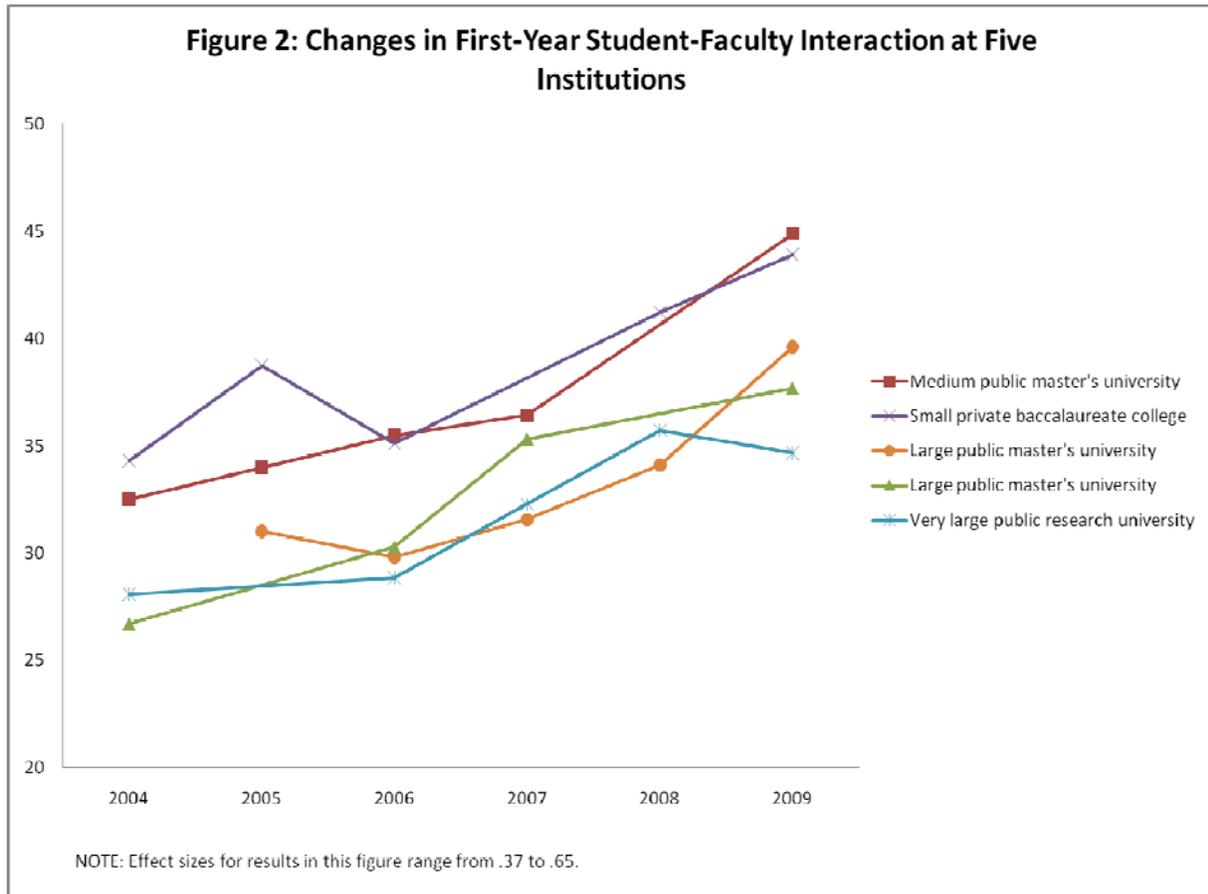
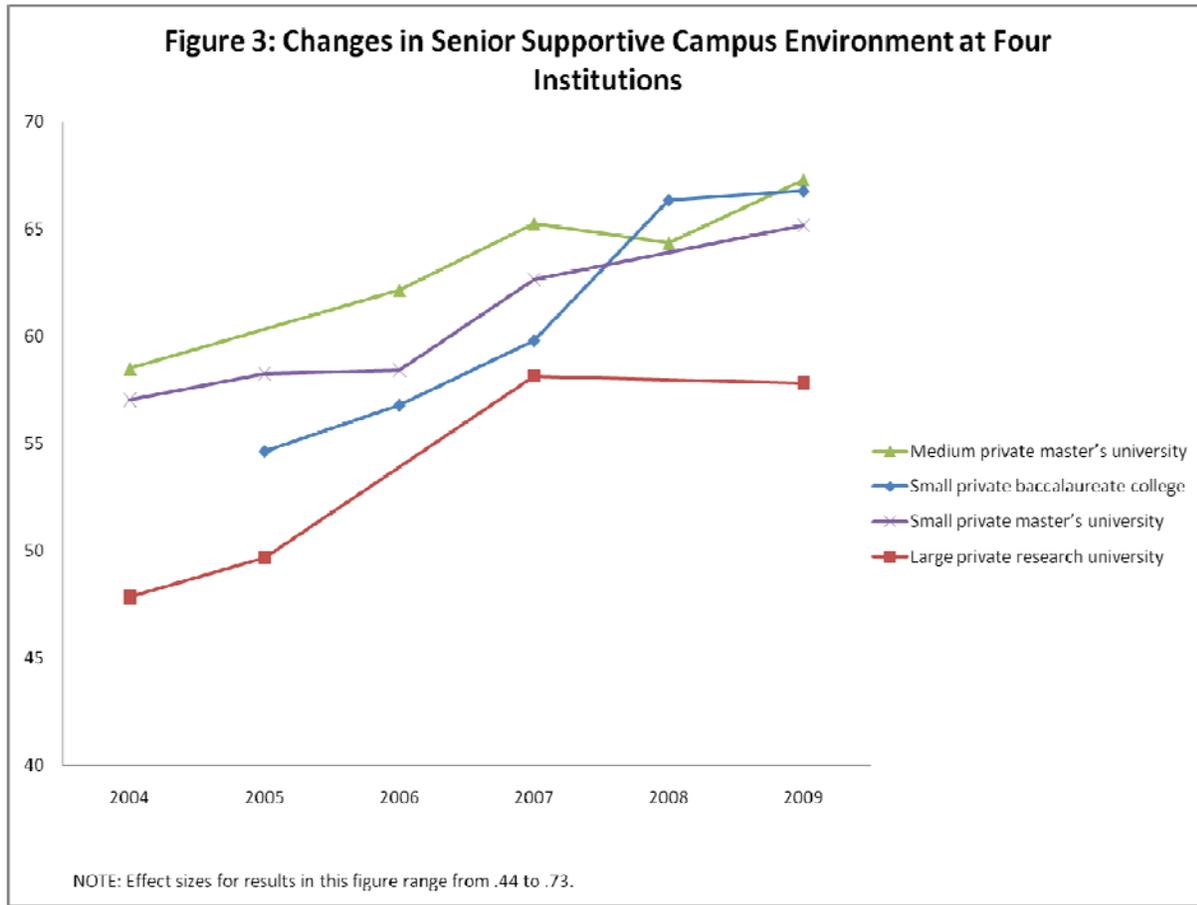


Figure 3.



Appendix A. Scales used for Analyses

Academic Challenge (First-Year Cronbach's Alpha = .71; Senior Cronbach's Alpha = .73)

- Number of hours in a typical week preparing for class (studying, reading, writing, rehearsing, and other academic activities)
- Number of assigned textbooks, books, or book-length packs of course readings during the current school year
- Number of written papers or reports of 20 pages or more during the current school year
- Number of written papers of between 5 and 19 pages during the current school year
- Extent to which coursework emphasized this school year: Analyzing the basic elements of an idea, experience or theory such as examining a particular case or situation in depth and considering its components
- Extent to which coursework emphasized this school year: Synthesizing and organizing ideas, information, or experiences into new, more complex interpretations and relationships
- Extent to which coursework emphasized this school year: Making judgments about the value of information, arguments, or methods such as examining how others have gathered and interpreted data and assessing the soundness of their conclusions
- Extent to which coursework emphasized this school year: Applying theories or concepts to practical problems or in new situations
- How often worked harder than you thought you could to meet an instructor's standards or expectations at your institution this school year?
- Extent to which your institution emphasizes: Spending significant amounts of time studying and on academic work

Active and Collaborative Learning (First-Year Cronbach's Alpha = .62; Senior Cronbach's Alpha = .63)

- How often asked questions in class or contributed to class discussions at your institution during the current school year?
- How often made a class presentation at your institution during the current school year?
- How often worked with other students on projects during class at your institution during the current school year?
- How often worked with classmates outside of class to prepare class assignments at your institution during the current school year?
- How often tutored or taught other students at your institution during the current school year?
- How often participated in a community based project as part of a regular course at your institution during the current school year?
- How often discussed ideas from your reading or classes with others outside of class (students, family members, coworkers, etc.) at your institution during the current school year?

Student-Faculty Interaction (First-Year Cronbach's Alpha = .70; Senior Cronbach's Alpha = .75)

- How often discussed grades or assignments with an instructor at your institution during the current school year?
- How often talked about career plans with a faculty member or advisor at your institution during the current school year?
- How often discussed ideas from your reading or classes with faculty members outside of class at your institution during the current school year?
- How often worked with faculty members on activities other than coursework at your institution during the current school year?
- How often received prompt feedback from faculty on your academic performance at your institution during the current school year?
- Done or plan to work with a faculty member on a research project

Supportive Campus Environment (First-Year Cronbach's Alpha = .77; Senior Cronbach's Alpha = .76)

- Extent to which your college emphasized this school year: Providing the support you need to succeed academically
- Extent to which your college emphasized this school year: Helping you cope with your nonacademic responsibilities (family, work, etc.)
- Extent to which your college emphasized this school year: Providing the support your need to thrive socially
- Rate the quality of relationships with other students
- Rate the quality of relationships with faculty members
- Rate the quality of relationships with administrative personnel and offices

High-Impact Practices

- Proportion of first-year students who participated in a learning community or in service-learning as part of a regular course,
- Proportion of seniors who participated in (a) a practicum, co-op, internship, or field experience, (b) research with a faculty member, (c) study abroad, (d) a culminating senior experience, or (e) service-learning.