9 Principles of Good Practice for Assessing Student Learning

1.The assessment of student learning begins with educational values. Assessment is not an end in itself but a vehicle for educational improvement. Its effective practice, then, begins with and enacts a vision of the kinds of learning we most value for students and strive to help them achieve. Educational values should drive not only what we choose to assess but also how we do so. Where questions about educational mission and values are skipped over, assessment threatens to be an exercise in measuring what's easy, rather than a process of improving what we really care about.

2.Assessment is most effective when it reflects an understanding of learning as multidimensional, integrated, and revealed in performance over time. Learning is a complex process. It entails not only what students know but what they can do with what they know; it involves not only knowledge and abilities but values, attitudes, and habits of mind that affect both academic success and performance beyond the classroom. Assessment should reflect these understandings by employing a diverse array of methods, including those that call for actual performance, using them over time so as to reveal change, growth, and increasing degrees of integration. Such an approach aims for a more complete and accurate picture of learning, and therefore firmer bases for improving our students' educational experience.

3.Assessment works best when the programs it seeks to improve have clear, explicitly stated purposes. Assessment is a goal-oriented process. It entails comparing educational performance with educational purposes and expectations -- those derived from the institution's mission, from faculty intentions in program and course design, and from knowledge of students' own goals. Where program purposes lack specificity or agreement, assessment as a process pushes a campus toward clarity about where to aim and what standards to apply; assessment also prompts attention to where and how program goals will be taught and learned. Clear, shared, implementable goals are the cornerstone for assessment that is focused and useful.

4.Assessment requires attention to outcomes but also and equally to the experiences that lead to those outcomes. Information about outcomes is of high importance; where students "end up" matters greatly. But to improve outcomes, we need to know about student experience along the way -- about the curricula, teaching, and kind of student effort that lead to particular outcomes. Assessment can help us understand which students learn best under what conditions; with such knowledge comes the capacity to improve the whole of their learning.

5.Assessment works best when it is ongoing not episodic. Assessment is a process whose power is cumulative. Though isolated, "one-shot" assessment can be better than none, improvement is best fostered when assessment entails a linked series of activities undertaken over time. This may mean tracking the process of individual students, or of cohorts of students; it may mean collecting the same examples of student performance or using the same instrument semester after semester. The point is to monitor progress toward intended goals in a spirit of continuous improvement. Along the way, the assessment process itself should be evaluated and refined in light of emerging insights.

6.Assessment fosters wider improvement when representatives from across the educational community are involved. Student learning is a campus-wide responsibility, and assessment is a way of enacting that responsibility. Thus, while assessment efforts may start small, the aim over time is to involve people from across the educational community. Faculty play an especially important role, but assessment's questions can't be fully addressed without participation by student-affairs educators, librarians, administrators, and students. Assessment may also involve individuals from beyond the campus (alumni/ae, trustees, employers) whose experience can enrich the sense of appropriate aims and standards for learning. Thus understood, assessment is not a task for small groups of experts but a collaborative activity; its aim is wider, better-informed attention to student learning by all parties with a stake in its improvement.

7.Assessment makes a difference when it begins with issues of use and illuminates questions that people really care about. Assessment recognizes the value of information in the process of improvement. But to be useful, information must be connected to issues or questions that people really care about. This implies assessment approaches that produce evidence that relevant parties will find credible, suggestive, and applicable to decisions that need to be made. It means thinking in advance about how the information will be used, and by whom. The point of assessment is not to gather data and return "results"; it is a process that starts with the questions of decision-makers, that involves them in the gathering and interpreting of data, and that informs and helps guide continuous improvement.

8.Assessment is most likely to lead to improvement when it is part of a larger set of conditions that promote change. Assessment alone changes little. Its greatest contribution comes on campuses where the quality of teaching and learning is visibly valued and worked at. On such campuses, the push to improve educational performance is a visible and primary goal of leadership; improving the quality of undergraduate education is central to the institution's planning, budgeting, and personnel decisions. On such campuses, information about learning outcomes is seen as an integral part of decision making, and avidly sought.

9.Through assessment, educators meet responsibilities to students and to the public. There is a compelling public stake in education. As educators, we have a responsibility to the publics that support or depend on us to provide information about the ways in which our students meet goals and expectations. But that responsibility goes beyond the reporting of such information; our deeper obligation -- to ourselves, our students, and society -- is to improve. Those to whom educators are accountable have a corresponding obligation to support such attempts at improvement.

Authors: Alexander W. Astin; Trudy W. Banta; K. Patricia Cross; Elaine El-Khawas; Peter T. Ewell; Pat Hutchings; Theodore J. Marchese; Kay M. McClenney; Marcia Mentkowski; Margaret A. Miller; E. Thomas Moran; Barbara D. Wright

Qualities of Good Assessment

Effective assessment aspires to inform change or confirm existing practice. Successful assessment yields <u>action</u> either by guiding change or affirming the continuation of current educational strategies.

High quality assessment. . .

1. *Focuses on what matters most.* There are always numerous "interesting questions" but assessment is not simply an intellectual exercise. High quality assessment springs from an organization's well defined goals, objectives, and mission. Clarity about desired outcomes is the first step in assessment.

2. *Focuses on elements that the organization can change.* Little can be gained by studying fixed characteristics or conditions which are exceedingly difficult to change. Work first to understand the impact of elements which the organization can change if it wishes to do so.

3. *Is built on the goodwill of participants and stakeholders.* Assessment efforts that are properly timed, orchestrated, and explained to participants are more likely to produce trustworthy data and outcomes.

4. *Is a means, not an end.* The purpose of assessment is to confirm or improve practice so as to create a desired end. The generation of raw data or assessment findings is but a step toward the desired end of creating information that guides future actions.

5. *Is multidimensional.* Assessment seldom meets the rigorous standards of controlled experimentation. Assessment is more credible when information is established through corroborating studies using multiple measures and methods.

6. *Includes input from all stakeholders.* Externally mandated and contrived assessment is less likely to produce meaningful change than assessment created by and for those directly involved in the assessed activity.

7. *Places findings in an appropriate context.* Comparative benchmarks, longitudinal data, or professional judgments are needed to provide context for assessment findings.

8. *Appropriately disseminates findings.* Assessment reports are 1) disseminated to individuals who shape the desired outcome, and 2) written in accessible language for each target audience. A general rule of thumb is that aggregated data may be widely disseminated - but data which identifies a specific individual should be treated with the highest level of confidentiality and disseminated with great care. Reports should be written at the appropriate level of specificity for each target audience (e.g. executive summary, concise edition, full report).
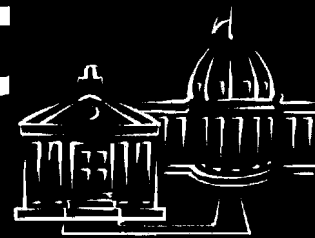
What's So Special About Assessment in the First Year of College?

Randy L. Swing

A STANDARD METAPHOR FOR THE FIRST COLLEGE YEAR IS AN OPEN DOOR welcoming new students as they move from high school, work, military service, or a "gap year." Humans have long acknowledged the importance of beginnings and transitions such as the first year of college. In Roman culture, Janus was believed to rule over doors and gates as the god of beginnings, and American folk wisdom acknowledges the importance of "starting right" as the first step in ensuring successful outcomes. Educators have assigned special importance to the first year of college as the foundation for the college experience, an element too important to leave to chance. Over the past two decades, college faculty and staff have developed a host of educational initiatives to undergird this transition period. First-year seminars are one such innovation and, according to a 2001 survey by the Policy Center on the First Year of College, can be found at 94 percent of accredited colleges and universities. Other common components of first-year programs include mandatory academic advising, course placement testing, early warning systems for low academic performance, learning community structures, and special residential programming.

Until recently, assessment of first-year activities was often limited to evaluating their impact on first-to-second-year persistence in college. Over the past five years, assessment of first-year students and structures has developed rapidly, spurred by the accountability agenda of The Pew Charitable Trusts. New tools and techniques have been developed so that two-year and four-year institutions can assess what students do during the first year of college and how they change as a result of participating in educational activities. These new tools join the handful of existing survey instruments that were specially designed to assess student experiences in their first college year (see Exhibit 1).

A plethora of research on what matters in the first year and an array of new outcome assessment tools have made possible the development of comprehensive assessment

Note: I am grateful to Randy L. Swing and colleagues for providing four articles for this special issue on assessment of the first year. —Editor

CONTENTS

strategies to measure learning outcomes and program-level impact of experiences during the first college year. It is now possible to consider the first college year as a discrete unit of analysis—a distinct time period intentionally designed to produce specific outcomes. As such, measurements of first-year students and programs can be viewed as summative assessments of this discrete set of educational experiences, and educators can establish how well they are meeting the goals established for the first year. First-year assessment can also be viewed as formative in that it establishes baseline and midprocess measurements early enough to allow faculty to change the learning outcomes desired by the end of a two-year or four-year degree program.

Summative measures of learning collected at the conclusion of an associate or bachelor's degree program are insufficient to help educators and students maximize learning and personal development across the college years. Managing the desired outcomes of college requires an understanding of when learning and growth occur during the educational process and which educationally purposeful initiatives or experiences are associated with gains for various subpopulations of students. Producing this kind of rich assessment of the first year is dependent on general principles of good assessment in higher education, such as ensuring that instruments are valid, reliable, and measure important constructs; that processes are inclusive of multiple constituencies; and that student time is respectfully and efficiently used. But assessment of first-year students and programs also presents unique challenges and opportunities that must be considered in addition to general good practices in assessment.

Collection of Baseline Data

A perennial tension in the first year is finding time to measure student skills, abilities, and cognitive knowledge at the point of entry without interfering with other goals for orienta-

Call for Contributions

The editor welcomes short articles and news items for *Assessment Update*. Guidelines follow for those who would like to contribute articles on outcomes assessment in higher education.

- *Content:* Please send an account of your experience with assessment in higher education. Include concrete examples of practice and results.
- *Audience: Assessment Update* readers are academic administrators, campus assessment practitioners, institutional researchers, and faculty from a variety of fields. All types of institutions are represented in the readership.
- *Style:* A report, essay, news story, or letter to the editor would be welcome. Limited references can be printed; however, extensive tables cannot be included.
- *Format:* In addition to standard manuscripts, news may be contributed via letter, telephone, or fax (317) 274-4651. The standard manuscript format is a 60-space line with 25 lines per page. If word processing is used, please submit a 3½" diskette and three paper copies of your article. Word is preferred. Articles may also be sent to <kblack@iupui.edu> as a Microsoft Word attachment.
- *Length:* Articles should be four to eight typed, double-spaced pages (1,000–2,000 words). Short news items and content for the Calendar and Book Review sections may be 100–500 words in length. Annotations of recent publications for the Resources feature should be about 50–100 words long.
- *Copyright:* Articles shall not have been registered for copyright or published elsewhere prior to publication in *Assessment Update*.
- *Deadlines:* Each issue is typically planned four months before its publication. Future deadlines for submitting articles are June 1 (September–October 2004 issue), August 1 (November–December 2004 issue), and October 1 (January–February 2005 issue).

Please address mailed contributions and comments to Trudy W. Banta, Editor, *Assessment Update*, Rm. 140 Administration Bldg., 355 N. Lansing St., Indianapolis, IN 46202-2896. ■

College Student Experiences Questionnaire & College Student Expectations
  Questionnaire
Center for Postsecondary Research, Indiana University Bloomington
<http://www.indiana.edu/~cseq/>

College Student Inventory
Noel–Levitz
<http://www.noel-levitz.com>

CIRP—The Freshman Survey & Your First College Year
Higher Education Research Institute, University of California, Los Angeles
<http://www.gseis.ucla.edu/heri/>

The College Student Report—National Survey of Student Engagement
Center for Postsecondary Research, Indiana University Bloomington
<http://www.indiana.edu/~nsse/>

Community College Survey of Student Engagement
Community College Leadership Program, University of Texas at Austin
<http://www.ccsse.org>

Community College Student Experiences Questionnaire
Center for the Study of Higher Education, University of Memphis
<http://www.people.Memphis.edu/~coe_cshe/ccseq_main.htm>

First-Year Initiative
Educational Benchmarking, Inc.
<http://www.webebi.com>

tion and the first days of college. Whether an institution is dealing with traditional-age residential students or commuting nontraditional students, there are multiple demands, questions, and details that need attention in the first days of college. Spending time on assessment does not rank high on the priority list of most new students, nor of student life professionals who are assigned to assist students during this time of transition. Still, the first days of college provide an excellent opportunity to capture baseline information that will be needed at a later point to answer the question, "Did our educational practices make a difference?" First-year assessment must be efficient, limiting data collection to essential information that can be used in future assessment efforts and avoiding duplication of information already available elsewhere.

It is imperative that tangible results come from the first experiences students have with collegiate assessment, to establish the importance of their participation in these and future assessment activities. Some campuses provide individualized student reports, sharing them directly with students or through academic advisers; others publicize results in the campus newspaper or in newsletters mailed to students and parents. Closing the loop in assessment includes feedback to participants; this is especially important in the case of first-year students, who will be asked to participate in future assessment efforts.

Age and Maturity Issues

Institutional review boards (IRBs) with oversight of human subject protection often require special handling for students who are under eighteen years of age. Some institutions require parental permission for underage students to par-

ticipate in assessments that ask individuals to report personal opinions or behaviors. It is possible, though time-consuming, to collect parental permission, but simply removing the small number of students who are underage from the survey population may have little or no impact on institution-level results. Assessment leaders should determine how many new students are under the age of eighteen (birth dates are often collected on admissions applications, so a simple computation can create a list of underage students for any targeted survey date), then consult their campus IRB early in the planning process to determine whether special actions should be considered for underage students.

Most first-year students, well trained by experiences in K–12 testing situations, take assessment tests seriously, but a small number of first-year students will exhibit immature behavior. The class clown who decides that his or her funny comments need to be shared with everyone in the testing room can be annoying and may inappropriately influence other students' responses. Assessment leaders should provide adequate supervision in large testing rooms or use other strategies to ensure that each survey is completed independently. Even in well-managed situations, a few students may rebel by providing extreme responses on multiple-choice items or using shock-value words in qualitative assessments. Since such responses fail to provide helpful information and can be misleading, cleaning data and isolating outlier and inappropriate responses should be a standard process in the analysis of data.

Another maturity issue may arise from the assessment professional. A common mistake in reporting first-year assessment outcomes is assigning all observed change to the investigated educational experiences, without acknowledging, especially for traditional-age students, that some change may result from the natural processes of maturing or adapting to a new situation. The use of control groups can allow alternative hypotheses to be tested,

but often it is not possible or practical to establish a research protocol that absolutely isolates cause and effect variables. In such cases, reports of first-year assessment initiatives should include appropriate statements about the limitations of the study.

It is challenging to use the same assessment instrument for traditional-age and nontraditional-age new students, whether one is measuring behaviors, opinions, or knowledge. Nontraditional-age students might, for example, find a knowledge test of mathematics intimidating if they have not been in a math class for some years. Likewise, they might find questions about friendships, dating, roommates, and alcohol use to be inappropriate for their own life situation. Students know when an instrument doesn't really make sense, and many will tire quickly when none of the response alternatives is exactly right for their situation. Assessment leaders must consider whether an assessment instrument will work with various subgroups of new students. The tools that an institution selects send messages to new students who are trying to figure out how they fit in at college.

Power Issues

First-year students know that they are novices in higher education, so they may not be comfortable criticizing teachers or staff members or be willing to report socially unpopular opinions. Assessment leaders must weigh the advantages of using student identifiers against the potential for bias that exists when students know their answers are not anonymous. In general, it is best to avoid heavy-handed assessment mandates or other power plays that force students to participate in assessment activities.

Power issues arise when assessments are mandated or conducted in captive audience settings. It may be true that "first-year students will do what you tell them to do," but this may not be true when posttest data are collected at a later point. It may be a mistake to compare results

from high-stakes testing, such as placement test scores, with later tests that have no consequences for the test taker.

One way to avoid power plays is to conduct assessments at the convenience of students. A focus group conducted at 8:00 A.M. is a bad idea when working with traditional-age first-year students; a better time may be 11:00 P.M.

Educational Jargon

Eventually college students pick up the educational jargon of their campus and even embrace its acronyms, but first-year students may be confused by unfamiliar terms. While educators understand terms such as "learning community," "first-year seminar," "experiential education," "humanities," and "critical thinking," these may be undecipherable concepts for new students. First-year assessment instruments should avoid educational jargon or provide definitions and examples when needed, to ensure that participants understand questions and instructions.

New students tend to view the institution as a single unit rather than a collection of uniquely named divisions. For example, an intended assessment of academic advising might actually reflect student experiences in both advising and registration, because students experience these as one event. Likewise, new students will understand that the teacher leads the class, but they may not be able to differentiate the various classifications of teachers, such as full professor, part-time instructor, lecturer, or graduate assistant. Instrument pilot testing is one way to locate and remove jargon and distinctions that new students may not understand.

Establishing Comparison Groups

Many first-year initiatives are mandated for all students regardless of major. Evaluating the impact of a campuswide mandatory program is problematic due to the lack of a control group of students for comparison. One solution is a cross-

sectional comparison using students from the year before the mandatory program was developed. There are inherent problems with cross-sectional comparisons; the most frequent mistakes are underestimating the halo effect of a new program and failing to consider that new programs seldom are perfect in their inaugural year.

Programs that are elective rather than mandatory present a different problem in establishing comparison groups. While there are "treatment" and "nontreatment" groups, the treatment group is different from the nontreatment group because participants volunteered. Ways to control for potentially confounding elements of student motivation, skills, experiences, and personalities must be included in assessment plans for elective programs.

Conclusion

The hypothesis of this essay is simply that assessing first-year students and programs is grounded in good practices in assessment but requires attention to uniquely significant conditions of the first year. Summoning the assistance of Janus might be helpful, but it is imperative that first-year assessment plans combine the knowledge of institutional researchers, assessment practitioners, and frontline faculty and staff who regularly interact with first-year students. Understanding the unique aspects of the first college year provides the necessary foundation for successful assessment activities, a means to the desirable goal of maximizing student learning and growth during the critically important first year. New students deserve no less than the best foundation experiences that colleges and universities can provide. ∎

Randy L. Swing is codirector and senior scholar at the Policy Center on the First Year of College located on the campus of Brevard College in Brevard, North Carolina.

Comparing Outcomes With and Without a Control Group

An early step in evaluating the learning outcomes of a first-year intervention (FYS, Learning Community, core curriculum course, etc.) is to determine if a non-treatment group exists for comparison of entering characteristics and later outcomes.

Model 1: An Experimental Design with random assignment to treatment or non treatment groups. (Rarely possible in educational interventions.) Differences in outcomes can be directly attributed to level of participation in the intervention, because all other student characteristics and experiences vary randomly.

```
        ┌──────────┐   ┌──────────────┐   ┌──────────────┐
        │          │   │     T₁       │   │  Learning/   │
 All 1ˢᵗ-│ Random   │──▶│  treatment   │──▶│  Behavioral  │
 year    │Assignment│   └──────────────┘   │  Outcomes    │
 students│          │                      └──────────────┘
        │          │         T₂
        │          │   ┌──────────────┐   ┌──────────────┐
        │          │──▶│ no treatment,│──▶│  Learning/   │
        │          │   │ shared other │   │  Behavioral  │
        └──────────┘   └──────────────┘   │  Outcomes    │
                                          └──────────────┘
```
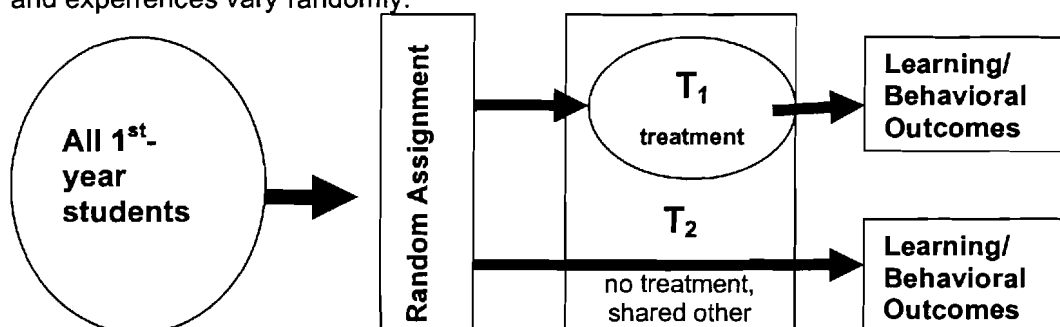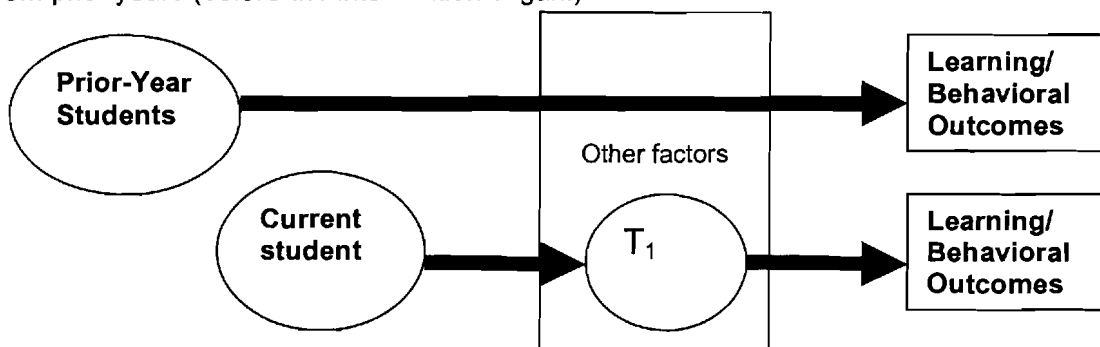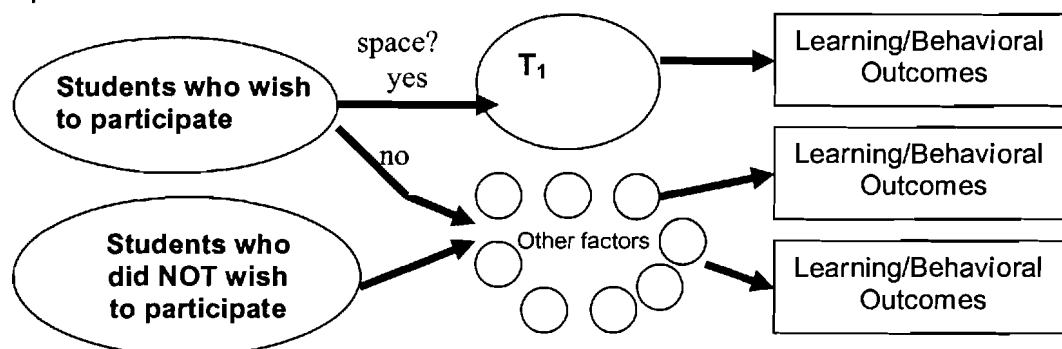
Model 2: Mandatory intervention models do not have contemporary comparison groups. They must use benchmarking (comparison to a similar external group) or be compared to students from prior years (before the intervention began.)

```
 Prior-Year                              ┌──────────────┐
 Students  ───────────────────────────▶  │  Learning/   │
                                         │  Behavioral  │
                          Other factors  │  Outcomes    │
                                         └──────────────┘
 Current                                 ┌──────────────┐
 student  ──────────▶  T₁  ───────────▶  │  Learning/   │
                                         │  Behavioral  │
                                         │  Outcomes    │
                                         └──────────────┘
```

Model 3: An elective intervention which does not fully meet the demand for enrollment produces three groups, 1) enrolled students, 2) students who wanted to but we denied enrollment, and 3) students who did not wish to enroll. Group 2 is a perfect control group for Group 1.

```
                    space?
 Students who wish   yes    T₁   ─────▶  ┌────────────────────┐
 to participate    ───────              │ Learning/Behavioral │
                        ╲ no            │     Outcomes        │
                         ╲              └────────────────────┘
                          ╲   ○ ○ ○  ─▶ ┌────────────────────┐
 Students who              ╲            │ Learning/Behavioral │
 did NOT wish   ──────────▶ Other factors○ Outcomes          │
 to participate                         └────────────────────┘
                          ○ ○  ○  ─────▶┌────────────────────┐
                                        │ Learning/Behavioral │
                                        │     Outcomes        │
                                        └────────────────────┘
```

Randy L. Swing, 2003. Policy Center on the First Year of College. swingrl@brevard.edu

Organizing First Year Assessment

<u>The good will of students</u> is the foundational premise undergirding any successful assessment effort. Bombarding students with surveys and tests is likely to discourage student participation, reduce the level of effort given, and spoil reliability of the measures. Assessment designs which judiciously use student time are more likely to produce appropriately high student effort and thus produce trustworthy data. *Institutions are responsible for building and maintaining the good will of students toward assessment.*

1. Use Sampling.
There is often little to gain by surveying every student. A carefully controlled sampling procedure will produce the same results as a population study. A limitation is that samples can not always be disaggregated to study every sub-population (the numbers get too small when the sample is divided into a lot of sub groups).

2. Look for Connections – Remember the Forest and the Trees.
Use surveys that cross administrative divisions. Avoid artificial divisions of the first year. Students see us as "one" institution, not a federation of units.

3. Divide and Conquer.
Use multiple opportunities for data collection. For example, collect one survey during summer orientation, another during move-in day, another at the first class, etc.

4. Use Existing Data Whenever Possible.
A data audit often reveals that institutions hold great quantities of useable data which can be connected in a central dataset. Whenever possible use student identifiers to link datasets – especially to make use of demographic data already on file. Many assessment efforts start with time-consuming data collection and bog down during the data analysis phase. Existing data reduces data collection time and increases time available to disaggregate, analyze, and report findings.

5. Embed Assessment in Courses.
When "regular course exams and assignments can also be used of assessment, student time is used wisely and many motivation issues are resolved. A limitation is that courses are not normally "randomly assigned."

6. Use Assessment Days.
Dedicate a day or half-day for assessment activities separate from class and orientation activities. Link assessment activities to a reward such as participation in early registration for the next academic term. James Madison University and Appalachian State University are model programs.

7. Sample Classes/ Sample Students
An alternative to sampling the student population is to sample the population of classes. A random or stratified sample of courses/class times will capture an appropriate sample of students. Have a team of peer assistants who visit the selected classes to administer the surveys. Clearly this requires support from faculty and advance planning.

8. Use Cohort-Based Assessment Cycles.
Consider every year's entering class to be a cohort that will be surveyed on one topic. Focus the assessment effort on topic X for the entering class of 2002, and then on topic Y for the entering class of 2003, etc. Over time, each cohort adds to the institutional knowledge base. This concept reduces the data overload often produced by assessment efforts. It may, however, result in some limitations due to cross-sectional methodologies.

9. Response Rates Matter.
Assessment built on low response rates is ripe for criticism and dismissal of the findings. Use a survey size and method that you can really control, follow-up, and "work" to accomplish response rates above 50% of the sample.

10. Response Rates DON'T Matter.
Are the respondents <u>representative</u> of the total population? A high response rate can still contain a systematic bias. Low response rates can be "random" and highly representative. Look for evidence that the sample is representative on gender, race, and other characteristics.

11. Involve Students.
Use focus groups, have students proctor surveys, involve students in interpreting results, and let students know how the institution uses assessment outcomes.

12. Signal the Importance of Assessment.
Don't assume that students or faculty know why the institution is conducting assessment. Use the appropriate mix of seriousness and fun, hold students accountable, intentionally establish a culture of assessment.