Design Issues to be Addressed in Proposed Student Growth Assessment:

1. **Assessment Objectives.** Are the student outcomes desired clearly stated? Exactly what knowledge, skills, or attitudes do we want students to develop as the result of a Purdue education? Do the specified outcomes apply to all Purdue majors-- or should some majors have different student outcome objectives than others? Specification of the desired learning outcomes is probably the most important issue in a student assessment system.

2. **Measures Validity.** Are the assessment measures selected a good match with the specified student outcomes (content validity)? Do the items on the assessment measures reflect the learning objectives stated in the assessment objectives? Do the selected measures have an established (published) track record that shows they actually measure what they are intended to measure (criterion validity)? Is there evidence that data produced by these measures either correspond well with other measures with established validity, or that they predict the kinds of future student outcomes expected? If the assessment measures have multiple scales or subscales (i.e., if one test measures a number of different capacities or qualities, do the test items within each subscale compose a unitary, coherent group (internal consistency)? Will students be motivated to give their true and best performance on these assessments, both in the freshman year and at graduation? With all the assessments they are already required to complete for grades in their academic courses, plus these additional University assessments, will they feel over-loaded with testing and thus not be motivated to perform well? This issue may be especially salient for graduating seniors. How would you get them to attend and complete the post-assessments with care, at a time when they are occupied with completing final exams, commencement, and planning for the future, including new jobs and moving away from Purdue? What would it take to create the conditions under which students would give their best performance, on both the entry and exit assessments?

3. **Adequate sampling.** Depending on the objectives of the student assessment, are the students taking the tests truly representative of the Purdue student body at large? Will we be able to confidently generalize the results from a sample of student assessments to the entire student population at Purdue? To specific colleges or majors? (Note: The best way to get accurate results would be to test every student. But doing that may be cost-prohibitive, or not practical for many other reasons. If the assessments are given to subgroups of students for practical or ethical reasons, we need to assure that the samples are not a “select” or “biased” group of students, such that their results would be different in some ways from the overall results expected for the entire student population. The best way to select a representative sample would be to randomly select students to do the tests, probably stratifying the selected participants according to important characteristics like gender, ethnicity, college, etc. But even with this approach, you need to have sufficient numbers to have the statistical power required to make comparisons among groups. The required sample numbers can be determined by
statistical power calculations. This should be done for each subgroup sample (college, major, etc.) of Purdue students to be assessed.) **How will the student assessment program deal with sample attrition?** Some students who are assessed at entry will not complete their degrees in the required time. They will leave Purdue or change majors one or more times before completing. Losing those students in the graduation post-assessment may invalidate the overall Purdue results, because their data will be missing, and they will undoubtedly be different in important ways from students who stay with their declared major and finish in 4 years.

4. **Basis for Comparison.** When looking at results of student assessments, either at one point in time, or when examining changes from entry to graduation, how will we know what level of performance represents “success”? For any student outcome measured, what would constitute “value-added,” compared to an individual who did not receive a Purdue education? It is not sufficient to conclude a Purdue education is “successful” or “unsuccessful” based on an increase, no change, or decrease in Purdue students’ test scores over 4 years’ time, unless we know what the results would have been if the students had NOT attended Purdue. In order to draw valid conclusions, we must know how an identical (or at least very similar, via careful matching) sample of students would perform, over the same time period, had they not had a Purdue education, or had they attended a different post-secondary institution. What would be a reasonable comparison group for Purdue students’ pre- and post-assessments? Do the selected assessment tests have current published national norms? If so, is the population upon which those norms are based a reasonable comparison group for Purdue students?