Level 4. Assessing programs
Some formative but mostly summative questions guide assessment of programs (e.g., general education or a major).

Typical assessment questions at this level:
- Do the program’s courses, individually and collectively, contribute to its outcomes as planned?
- How well does the program fulfill its purposes in the entire curriculum?
- How well do the program’s sub-categories (e.g., distributive requirements in general education) contribute to the overall purposes?
- Does the program’s design resonate with its expected outcomes?
- Are the courses organized in a coherent manner to allow for cumulative learning?
- Does the program advance institution-wide goals as planned?

Sources of evidence: Direct evidence of student learning from many sources can contribute to program-level assessment: assignments from individual courses, student portfolios built over the program’s duration, entering student tests or assignments, capstone projects, results of common assignments, commercial tests. Selected assignments from other programs can be re-scored (given a “second reading”) by program faculty (e.g., to assess the general education program’s success in developing such institution-wide goals as communication, quantitative literacy, critical thinking, or ethical responsibility). Given the number of potential data sources and the amount of evidence that could be amassed, careful planning is needed to identify the important points for sampling and analysis. Program assessment may likely involve several sources of evidence gathered at the point of entry, a midpoint, and at the end of the program. End point data is particularly valuable as a summative indicator of how well the program, taken as a whole, is achieving its goals. Individual student grades are not informative at this level.

Aggregation of data: Course-level assessments of the courses in a program can be analyzed individually or collectively to reveal whether program goals are being achieved; sampling might be prudent in a large program. Information about the sub-categories in a program (e.g., distribution areas) can be aggregated to the program level (e.g., general education). Sampling of student portfolios considered excellent, average, and sub-par can vividly portray growth in student performance from beginning to the end of a program. Disaggregated data can reveal how sub-groups of students are succeeding in the program. Some external, commercially available assessments can be compared to norms (e.g., the Major Field Tests from ETS).

Data uses:
- To confirm the purpose of the program (e.g., its place in the entire curriculum or connection to mission)
- To check alignment of program design with program outcomes
- To discern how well the program, from its beginning to end, fosters cumulative learning of the desired outcomes
- To discover how well the program as a whole enables students to achieve end-point levels of competence for all program outcomes
- To identify superfluous and/or missing curricular and co-curricular elements in the program

Responsibilities: Responsibility largely rests on the program faculty, collectively and individually. Collectively, the faculty assumes responsibility for the entire program achieving its—and relevant institution-wide—goals and outcomes. Individual instructors are responsible for advancing the program and institutional goals embedded in their courses. Faculty members cooperate in establishing program “standards” and scoring rubrics for the quality of work expected.

Levels of Assessment: From the Student to the Institution, by Ross Miller and Andrea Leskes (2005)