Assessment Glossary

PART 1: Assessment & Strategic Planning

Assessment is the general term that describes the process of using evidence to learn if we are achieving our goals for learning, institutional effectiveness, etc. Henning (2006): "Assessment is helping folks determine if they are doing what they say they are doing and determine how well they are doing it so that they can make improvements in what they are doing."

The **Assessment Cycle** includes identifying outcomes, selecting an assessment method, aligning and delivering the learning opportunity/resource/service, administering the assessment, collecting and reporting the data, analyzing the results, and developing an action plan. Once the loop is completed, it is repeated to assess the impact of the action plan.

Assessment Methods are devices, tools, or mediums used to collect information/data. Examples are surveys, focus groups, document review, portfolio or capstone evaluation, interviews, observation, or other research methods.

An **Assessment Plan** articulates the actions one intends to take to complete an assessment project or set of projects. This includes department or program mission and goals, program or learning outcomes, assessment methods, data analysis strategies, and plans for interpreting and sharing results.

Baseline (Anthology/Campus Labs) **is the <u>only</u> survey software used by Student Affairs staff for measuring learning outcomes.** Baseline provides a centralized, accessible location for Division assessment activities and data. Baseline contains data collection and reporting tools for surveys, rubrics, and polling. Baseline

provides the ability to create dashboards that show progress over time and document key performance indicators. Staff can choose to use Qualtrics for gathering data to inform operations and/or programmatic outcomes.

Baseline Data is information that serves as a basis for comparison in assessing a program's impact or effectiveness.

Benchmarking is the process of utilizing data to compare the attributes or performance of a program, organization/institution or group with the attributes or performance of similar programs, peer organizations/institutions, or comparison groups. Benchmarking is useful in understanding larger trends in a field or discipline to make meaning of local data, set goals, and advance best practices.

Closing the Loop means using assessment findings to inform decisions to improve practice, and sharing findings with stakeholders.

A **Dashboard** is a summary view of key performance indicators relevant to a particular goal or objective.

A **Direct Measure** is a method of collecting information that requires students or other stakeholders to display their knowledge and skills in ways that are directly observable (Palomba & Banta, 1999). Learning outcomes are most appropriately measured with direct measures. (observing task performance, pre/post test, evaluating a portfolio, etc.)

A **Goal** is a desired end result written in broad terms.

Indirect Methods are methods of collecting information that measure participants' perceptions, perspectives, or opinions. (e.g. Likert survey items, focus groups)

An **Institutional Review Board (IRB)** is a group within an institution of higher education providing oversight to ensure the appropriate steps are taken to protect the rights and welfare of human subjects in research.

Interpreting Results refers to the process of making meaning of data that has been collected. The process should involve consideration and when possible discussion with key stakeholders to understand the significance of assessment results, the possible reasons for those results, and possible actions to address or improve the results.

A **Key Performance Indicator** is a quantifiable metric or data point used to evaluate performance in terms of advancing a strategic goal, objective, or strategy.

Learning Outcomes specify what a student or participant should know or be able to do as a result of engaging in a learning opportunity.

Longitudinal Assessment involves repeated measurement or observation of the same variables at more than one point in time.

Measures is a general term for the tools we use to collect data.

A **Mission** describes an organization's purpose. It provides a means of succinctly communicating what we do.

Mix Methods refers to collecting, analyzing, and "mixing" both quantitative and qualitative research and methods in a single study to understand an issue, experience, environment, etc.

Needs Assessment is the process of collecting and interpreting data to determine a campus, community, or organization/institution's needs. It is sometimes referred to as conducting an "environmental scan." Results can be leveraged to allocate resources, or structure programs, services, facilities, policies or processes to best meet the needs of stakeholders.

An **Objective,** like a goal, communicates the intended effect of a service or intervention. However, an objective is more specific than a goal and identifies areas of focus that can result in achievement of a broader goal.

Outcomes articulate the measurable expected results or impacts of an activity or program effort.

A **Program Outcome** specifies what a unit, program, or process intends to do or accomplish to improve operations or support achievement of institutional, divisional, or departmental goals and objectives or priorities.

Program Review is a process designed to enhance organizational performance via the systematic review of data pertaining to department activities, service delivery and use, resource management, and contributions to the advancement of the Student Affairs mission and the University's strategic plan. For most functional areas, Student Affairs employs CAS standards in structuring Program Review.

Qualitative Data is data that approximates or describes but does not apply numeric measurement to define the characteristics, or properties of an environment, experience, event, person, phenomena, situation, etc. Examples of qualitative data are gender identity, college major, or hometown.

Quantitative Data is data that can be meaningfully expressed as a number. Examples of quantitative data are GPA, service hours completed, or semesters enrolled full-time.

Research involves the collection and analysis of data to build or confirm a theory or hypothesis.

Strategic Planning is a process used to clarify an organization's purpose and goals, and identify strategies or actions to achieve specified goals.

Values serve as guiding principles for an organization.

Vision Statements are aspirational in nature. As such, a vision statement should reflect an organization's values and mission. They can act as a launching pad for strategic planning.

PART 2: Data Analysis & Statistics

Drawn from Auburn University Student Affairs glossary of assessment terms

Confidence Interval A range of values so defined that there is a specified probability that the value of a parameter lies within it.

Control Group The group in an experimental design that receives either no treatment or a different treatment from the experimental group. This group can thus be compared to the experimental group.

Correlation A common statistical analysis, usually abbreviated as "r", that measures the degree of relationship between pairs of interval variables in a sample. The range of correlation is from -1.00 to zero to +1.00. Also, a non-cause and effect relationship between two variables.

Crosstab A table providing the number and/or percentages of responses in a particular combination based on answers from two questions (e.g. Responses to a Likert Scaled question by gender identities).

Dependent Variable (DV) The variable that is expected to change when an independent variable is changed. Example: Researcher wants to know the impact of student involvement on grades. The dependent variable is the impact on grades.

Descriptive Statistics Descriptive statistics are the most common means of summarizing quantitative data. Descriptive statistics include measures of central tendency (mean, median, mode), measures of variation (standard deviation, variance), and relative position (quartiles, percentiles).

Distribution represents the variability and dispersion of scores in a data set.

Generalizability refers to the extent to which findings or results conducted on one sample can be applied to another sample or the population at large.

A **Hypothesis** is a tentative explanation based on theory to predict a relationship between variables.

An **Independent Variable (IV)** is a variable that is intentionally changed to observe its effect on another variable. Example: You want to know the impact of student engagement on sense of belonging. The independent variable is level of engagement.

The **Mean** is the average value for a set of responses. It is computed by adding all the values and dividing by the number of values.

The **Median** is the middle value of a set when the values that have been ordered by rank. For example, in this set of numbers (0, 0, 0, 1, 1, 2, 2, 2, 2, 3, 3) the value 2 is the Median value as 2 is the middle value with 5 values falling below it and 5 numbers falling above it. If there is an even number of values in the set, the median is the average of the two numbers in the middle.

The **Mode** is the most frequent value in a set of data.

A **Null Hypothesis (H0)** reflects that there will be no observed effect during statistical testing. For example, if you are interested in understanding whether sense of belonging differs across gender identities the Null Hypothesis would be that there is NO difference in levels of sense of belonging across gender identities. If your statistical test indicates there is a significant difference in sense of belonging associated with gender identity, you would reject the Null Hypothesis.

Population (N) An entire group of individuals as opposed to a sample (e.g. all people living in the state of NC, all undergraduates attending UNC, all 18 to 21 year old students attending 2 Year Colleges).

Random Sampling is a process used to draw a sample or subset of a population strictly by chance. Drawing participants by chance can reduce bias and increase the likelihood that findings are generalizable or somewhat representative of the larger population.

Reliability is the degree to which a measure yields consistent results.

A **Representative Sample** is a sample in which the participants closely match the characteristics of the population, and thus, all segments of the population are represented in the sample.

A **Sample (n)** is a subgroup of a population (e.g. 20% of all undergraduates attending UNC).

Standard Deviation is a measure of variation that indicates the typical or average distance between the scores in a set of responses and the mean score of the set of responses.

Standard Error is a statistical term that measures the accuracy with which a sample represents a population.

Statistical Significance is the probability that the difference between the scores of two or more groups are large enough that it is unlikely due solely to chance. Most practitioners agree that a significance value (P-value) of .05 or less [i.e., there is a 95% probability that the differences are real] sufficiently determines significance.

Statistical Tests are mathematical techniques used to test hypotheses to understand relationships between variables or trends in a data set. Examples are t-Tests, ANOVA's, and regression.

Validity refers to the degree to which a study, an instrument, an item or group of items accurately reflects or assesses the specific concept or concepts the practitioner intends to measure.