

## Assessment and Statistical Vocabulary

**Accountability** – The demand by a community (public officials, employers, and taxpayers) for school officials to prove that money invested in education has led to measurable learning.

**Action Plans** – The statement that indicates the specific changes that a given area plans to implement in the next cycle based on assessment results.

**Affective Outcomes** – Outcomes of education that reflect feelings more than understanding; likes, pleasures, ideals, dislikes, annoyances, values.

**Annual Report** – A report from each unit based on its goals and accomplishments as well as an assessment plan that is submitted annually. The report outlines how evidence was used to improve student learning outcomes and program outcomes through changes or to document that no changes were needed.

**Assessment** – The systematic collection, review, and use of information about educational programs undertaken for the purpose of improving student learning and development.

**Assessment Cycle** – The assessment cycle in higher education is generally annual and fits within the academic year. Outcomes, targets and assessment tools are established before the fall semester begins; data is collected by the end of spring semester; results are analyzed during the summer and early fall.

**Assessment tools (Methods)** - instruments used to gather data on student learning or program outcomes; either quantitative or qualitative selected in relation to the specific outcome(s) to be assessed; evidence available; learning to be observed and agreed-upon procedures for engaging in the assessment.

**Cohort** – A group whose progress is followed by means of measurements at different points in time.

**Continuous Variables** – Variables that can assume all values between any two given values - i.e. the time it takes for you to do your summer reading.

**Dependent Variable** – The variable on which we examine the impact of the manipulation of the independent variable. I.e. If we change the Independent variable, do we see a change in the dependent variable?

**Direct Assessment** – Assessment to gauge student achievement of learning outcomes directly from their work.

**Discrete Variables:** Variables that assume values that can be counted and will assume a succinct value. i.e. number of days it rained in Hawaii in 2009.

## **Empirical Data**

Qualitative Data - ways of collecting information that are concerned with understanding or conveying meanings or contexts, rather than making statistical inferences. Common forms: participant observations, focus groups, in-depth interviews, etc.

Quantitative Data - information that is collected or represented numerically; typically focuses on counting occurrences or measuring characteristics or behavior rather than meanings; easy to analyze statistically. Common forms: surveys, experiments, questionnaires, etc.

**Formative assessment** – The assessment of student or program achievement at different stages. The focus of formative assessment is on the documentation of student development or program improvement over time. It can be used to provide feedback, to modify, shape, or improve a program/service and/or student learning.

**Independent Variable** – A variable that is manipulated by an experimenter.

**Indirect Assessment** – Assessment that deduces results of outcomes through the reported perception of by students and other agents.

**Learning outcomes** – Operational statements describing specific student behaviors that evidence the acquisition of desired goals in knowledge, skills, abilities, capacities, attitudes or dispositions. Learning outcomes can be usefully thought of as behavioral criteria for determining whether students are achieving the educational goals of a program, and, ultimately, whether overall program goals are being successfully met. Outcomes are sometimes treated as synonymous with objectives, though objectives are usually more general statements of what students are expected to achieve in an academic program.

**Mean** - The most common expression for the mean of a statistical distribution with a discrete random variable is the mathematical average of all the terms. To calculate it, add up the values of all the terms and then divide by the number of terms. This expression is also called the arithmetic mean.

**Median** - The median of a distribution with a discrete random variable depends on whether the number of terms in the distribution is even or odd. If the number of terms is odd, then the median is the value of the term in the middle. This is the value such that the number of terms having values greater than or equal to it is the same as the number of terms having values less than or equal to it. If the number of terms is even, then the median is the average of the two terms in the middle, such that the number of terms having values greater than or equal to it is the same as the number of terms having values less than or equal to it.

**Mode** - The mode of a distribution with a discrete random variable is the value of the term that occurs the most often. It is not uncommon for a distribution with a discrete random variable to have more than one mode, especially if there are not many terms. This happens when two or

more terms occur with equal frequency, and more often than any of the others. A distribution with two modes is called bimodal. A distribution with three modes is called trimodal.

**Norm** – A distribution of scores obtained from a norm group. The norm is the midpoint (or median) of scores or performance of the students in that group. Fifty percent will score above and fifty percent below the norm.

**p-value:** “Probability” level. The likelihood that the difference observed between two interventions could have arisen by chance. It does not tell you the importance of the difference. The usual p value is arbitrarily set at 0.05.

**Performance-Based Assessment** – Direct, systematic observation and rating of student performance of an educational objective, often an ongoing observation over a period of time, and typically involving the creation of products. The assessment may be a continuing interaction and should ideally be part of the learning process. The assessment should be a real-world performance with relevance to the student and learning community. Assessment of the performance is done using a rubric, or analytic scoring guide to aid in objectivity. Performance-based assessment is a test of the ability to apply knowledge in a real life setting or performance of exemplary tasks in the demonstration of intellectual ability.

**Population** – All elements that are being studied comprise a population. I.e. If we want to study the distribution of IQ scores among politicians in Florida, the population will be the IQ scores. In statistics, the way we approach a problem and the formulas we use are completely dependent on whether we are examining a population or a sample.

**Portfolio** – A systematic and organized collection of a student's work that exhibits to others the direct evidence of a student's efforts, achievements, and progress over a period of time. The collection should involve the student in selection of its contents, and should include information about the performance criteria, the rubric or criteria for judging merit, and evidence of student self-reflection or evaluation.

**Random Sample** – A sample selected so that each group (of equal size) has the same chance of being selected.

**Reliability** – An assessment tool's consistency of results over time and with different samples of respondents.

**Rubric** – A set of criteria specifying the characteristics of a learning outcome and the levels of achievement in each characteristic.

**Sample** – A subset of the population. I.e., Listing the IQ scores of every tenth politician in Florida from an alphabetical listing of politicians. In statistics, the way we approach a problem and the formulas we use are completely dependent on whether we are examining a population or a sample.

**Statistical Power** - The power of a study is the ability for that study to find a difference between two treatments if the difference *really exists*. Power depends on the number of individuals in the study and the magnitude of the difference. A power of 0.80 is the standard. For example, a study of only 30 patients might not find a small difference between two drugs, whereas a study of 300 patients might find a difference.

**Summative assessment** – The assessment of student or program achievement at the end point of learning or a process. The focus of summative assessment is on the judgment of a current program, service, student learning, for quality and/or worth, based on previously established standards. It does not reveal the pathway of development to achieve that endpoint.

**Triangulation** – The collection of data via multiple methods in order to determine if the results show a consistent outcome.

**Validity** – The degree to which an assessment measures (a) what is intended, as opposed to (b) what is not intended, or (c) what is unsystematic or unstable