ANALYZING ASSESSMENT RESULTS

Once you have completed your assessment plan and have collected your data, you will need to analyze the results. This section introduces some basic methods of summarizing and presenting data. For a more detailed description on methods of analysis, refer to Linn & Miller (2005).

Scores

Analyzing data begins with scores. As mentioned in the previous section, the assessment methodology you choose will dictate what type of scores will be reported. For example, if you are using an objective assessment to measure student learning then your instrument will be scored dichotomously and summed across items to produce a total score. Test results may be presented in the form of the number or percent correct for each student.

If you are using standardized assessments, scores may be reported in the form of raw or standard scores for each student. Raw scores are the original and untransformed scores before any operation is performed on them. They are essentially meaningless by themselves but form the basis for other more interpretable scores such as percentiles, stanines and standard scores. They are usually reported as aggregates in the form of total scores and sub-scale scores.

Performance assessments require the use of a standardized scoring procedure usually involving a rubric. Results are usually presented in the form of scale scores for each student on each dimension. Dimensions are not usually summed since they generally represent distinct concepts such that a total score would be meaningless. The diagram on the following page from the University of Virginia's Assessment website illustrates how rubric results can be analyzed. In the diagram, students are represented in the first column of the table. Across the top of the table are dimensions labeled outcomes taken from a rubric. Each student is assigned a scale score ranging in value from 0 to 4 for each of the dimensions presented where 0 represents the lowest score and 4 the highest. From these data, descriptive statistics can be easily generated.
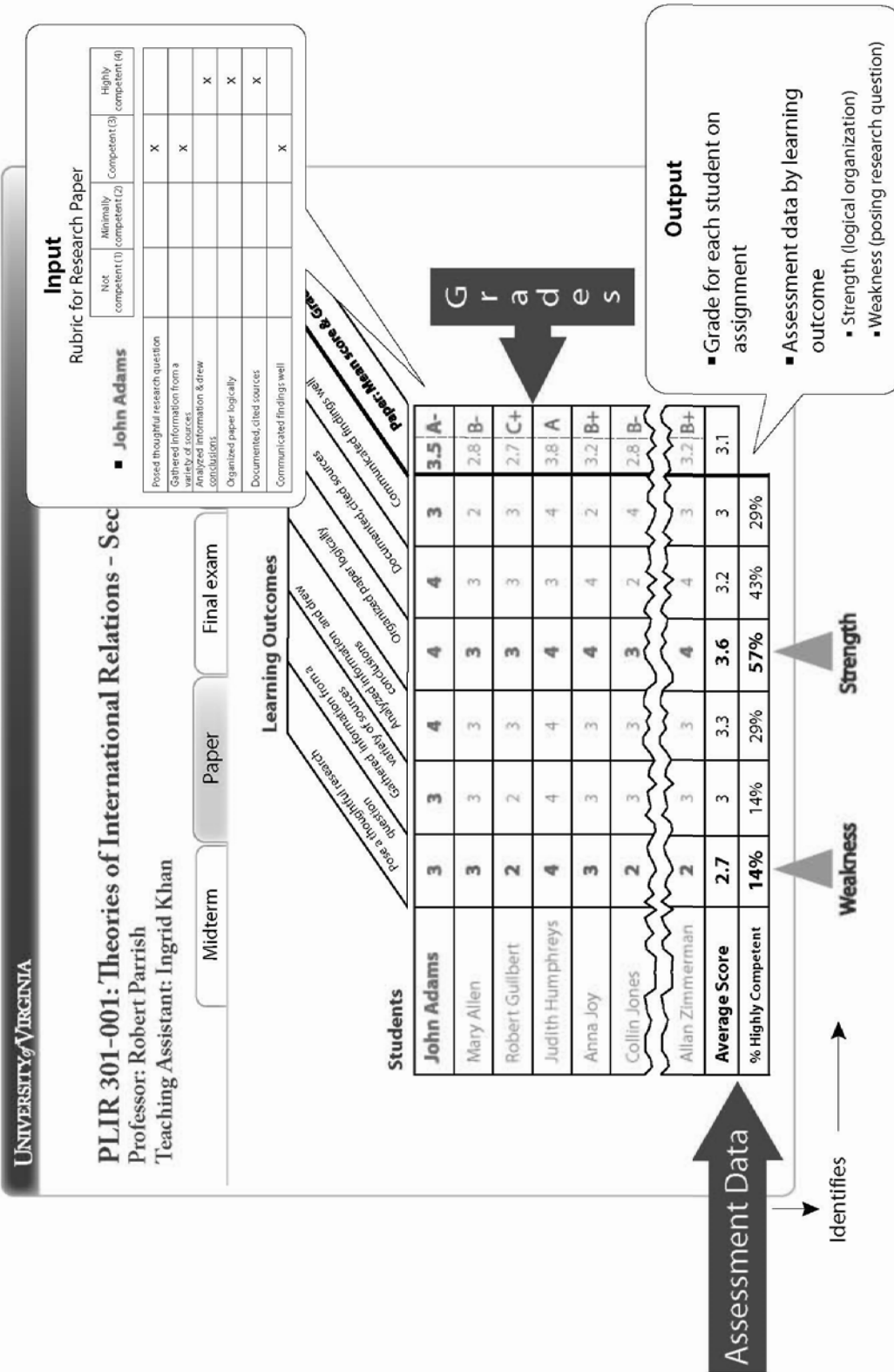
Once data are tabulated than they can be analyzed with the assistance of software packages available on campus such as EXCEL or SPSS.

Describing Quantitative Data

Once you have collected and scored your assessment data, you are ready to analyze and describe the results. Hidden among your scores there is an important message, possibly one that will help you make improvements in your program. Your main responsibility is to describe the data as clearly, completely, and concisely as possible. The statistics presented in this section will be helpful in this process.

One way of organizing data to get a clearer picture of what your scores mean is to compute frequencies. For example, you may be interested in determining the number of students who receive a particular score on an objective test or on a rubric. Frequency distributions are an easy way to summarize this information. Data are organized into classes of single values rather than grouped data and the number of occurrences for each single class of values can be reported. Frequencies are easily converted to percentages by dividing counts by the sum of all frequencies and then multiplying by 100. An easy way to display these data is in a table.

Using a Rubric to Produce Both Grades and Assessment Data

UNIVERSITY of VIRGINIA

PLIR 301-001: Theories of International Relations – Sec
Professor: Robert Parrish
Teaching Assistant: Ingrid Khan

Midterm	Paper	Final exam

Input

Rubric for Research Paper

■ John Adams

	Not competent (1)	Minimally competent (2)	Competent (3)	Highly competent (4)
Posed thoughtful research question			x	
Gathered information from a variety of sources			x	
Analyzed information & drew conclusions				x
Organized paper logically				x
Documented, cited sources				x
Communicated findings well			x	

Output

- Grade for each student on assignment

- Assessment data by learning outcome
  - Strength (logical organization)
  - Weakness (posing research question)

Grades

Learning Outcomes

Paper: Mean score & Grades

Students	Pose a thoughtful research question	Gathered information from a variety of sources	Analyzed information and drew conclusions	Organized paper logically	Documented, cited sources	Communicated findings well	Grade
John Adams	3	3	4	4	4	3	3.5 A-
Mary Allen	3	3	3	3	3	2	2.8 B-
Robert Guilbert	2	2	3	3	3	3	2.7 C+
Judith Humphreys	4	4	4	4	3	4	3.8 A
Anna Joy	3	3	4	4	4	2	3.2 B+
Collin Jones	2	3	3	3	2	4	2.8 B-
Allan Zimmerman	2	3	3	4	4	3	3.2 B+
Average Score	**2.7**	**3**	**3.3**	**3.6**	**3.2**	**3**	**3.1**
% Highly Competent	**14%**	**14%**	**29%**	**57%**	**43%**	**29%**	

Weakness

Strength

Assessment Data → Identifies

Calculating measures of central tendency and measures of dispersion are two other ways of summarizing data. Measures of central tendency describe the average or typical value of a set of scores. For example, you may be interested in determining the average score on an objective test or an average score for an outcome on a rubric. The three most commonly used measures of central tendency include the mean, median and mode. The mean is simply the arithmetic average that is obtained by summing all the scores in a set of data and then dividing by the number of scores (Mean = sum of all data/sample size). When calculating the mean, be careful of outlying data points that may artificially drag the mean up or down. In the case of extreme outlying points, it is better to rely on the median. The median is a counting average. If the number of scores is odd then the median is calculated by arranging the set of scores in ascending order and then counting up to the midpoint. If the number of scores is even, it is the average of the two middle scores. The mode is the most frequently occurring score. A set of scores sometimes has two modes (bimodal).

Measures of dispersion or variability describe how scores are spread out above and below the measures of central tendency. The range, variance, and standard deviation are examples of measures of variability. The range is the simplest measure of variability. It describes the difference between the two most extreme scores (range = maximum score – minimum score). The variance of your data is simply a measure of the average of your squared deviations from the mean. For each data point, the mean is subtracted from it and then this value is squared. The squared values are then added together and divided by the sample size minus 1 to create an average [variance = (sum of each data point - mean)^2/(sample size - 1)]. Finally, the standard deviation is the square root of the variance, and is in standard deviation units and not squared units to make interpretation easier [standard deviation = square root (variance)].

Drawing Conclusions about the Data

Now that you have analyzed the data, what do the results mean? The easiest place to start is with your learning outcomes. For each learning outcome, compare your results with the target of performance you set in your assessment plan and determine if your students met or failed to meet each target. Next, make sure for each outcome that the conclusion drawn is valid in light of your assessment methodology. For example, did the sample you selected reflect all of your students in the program in terms of student demographics (i.e., gender, grades, and/or class level, etc.)? Was the instrument you selected valid and reliable (see Section 4)? Did it do a good job discriminating between high and low scorers? Was scoring consistent across raters? Did results follow expected patterns? If your methodology is flawed, then it is important to interpret your results in light of these limitations. Finally, for each learning outcome, try to identify causes for success or failures within the program such as in the curriculum or in the academic process itself. Once causes have been identified it will be easy to devise the appropriate solutions for making improvements. These improvements might include changes to the program's assessment plan, changes to the curriculum, or changes to the academic process. The table on the following page details examples of changes that might be made as a result of assessment.

Examples of Changes that May be Implemented as a Result of Assessment

Changes to the Assessment Plan	• revision of intended learning outcomes • revision of measurement approaches • changes in data collection methods • changes in targets/standards • changes in the sampling
Changes to the Curriculum	• changes in teaching techniques • revision of prerequisites • revision of course sequence • revision of course content • addition of courses • deletion of courses
Changes to the Academic Process	• revision of admission criteria • revision of advising standards or processes • improvements in technology • changes in personnel • changes in frequency or scheduling of course offering

Adapted from University of Central Florida UCF Academic Program Assessment Handbook February 2005 Information, Analysis, and Assessment)

Closing the Loop

Closing the loop is the last phase in the assessment cycle and involves making decisions about how to respond to your program's shortcomings that have been identified through assessment data. Moreover, it is a dynamic process that involves shared feedback and collaborative reflection on the part of the faculty in the program. This begins first with making faculty aware of assessment findings and then organizing discussions around how to make improvements. Disseminating assessment findings is the first step. This may be accomplished through faculty news letters, informal emails, websites, and/or faculty meetings and retreats. Once this has been accomplished then faculty must decide what changes are needed and how they are going to make them. The most common types of changes often relate to the assessment plan, the program's curriculum and/or the academic process. When making plans for modifications, remember that changes should be manageable in terms of available time and resources. It is important not to make too many changes at once because it will be difficult to manage. Limit modifications to at most two per year depending on their magnitude. Finally, remember that improvements are generally gradual and cumulative in nature rather than all of a sudden, so don't get discouraged if they do not happen right away.