Interpreting Assessment Data

Assessment data is systematically gathered, analyzed and interpreted to understand unit and/or program functioning or effectiveness and to measure student learning within the context of a particular program or set of programs. The process of gathering and interpreting data helps us provide high quality programming, measure student learning, and communicate the ways Student Affairs programs and services contribute to meeting the academic mission of the University.

To develop an understanding of data analysis and interpretation, first we need a basic definition:

Data analysis and interpretation is the process of assigning meaning to collected information and determining the conclusions, significance, and implications of the findings (OIRA, Syracuse University).

For scholars conducting research in the field of Student Affairs, data interpretation and analysis often involves the application of advanced statistical techniques. This isn't necessarily the case for practitioners of Student Affairs conducting basic assessment. In this case, we need foundational knowledge and the ability to think critically about the information we collect. Before presenting on some of the types of assessment or data you might work with, it is helpful to recognize that there are a number of general frameworks or lenses you can apply to data analysis and interpretation. For example, you might be interested in exploring:

- **Differences** – Do students learn or develop more if they participate in a program compared to students who did not participate?
- **Relationships** – What is the relationship between participation and other outcomes?
- **Change** – Do students change over time as a result of participation?
- **Competency** – Do students achieve specified learning outcomes?

If the goal of interpreting data is to draw meaning from it, you need to know whether the information you have collected is reasonably accurate: Look through the data to discern whether there are any patterns that could mistakenly affect interpretation. For example, look for patterns of missing data (e.g., most participants skipped the same question), patterns of miscoded or impossible data, and outliers (e.g., You give a survey to 10 people and 9 offer responses that indicate very high satisfaction with the service but one rates the service with the lowest mark in every category. This person would represent a statistical outlier; whereas their opinions are important, you might not want to use a numeric average to interpret data as the very low ratings of one person might skew or misrepresent the overall nature of the data) or other irregularities (e.g., there were 25 questions on the survey and your data file contains information for 24 questions).

Once you have a sense that the data is relatively accurate, you can work on getting it into a more useful form. This usually involves the summary or reduction of individual data points. How we summarize data relates to the kind of data we have collected.

Quantitative and Qualitative Data

The Quantitative/Qualitative dichotomy represents a basic means of classifying or understanding data.

The term *quantitative data* is used to describe data that can be meaningfully expressed as a number, or quantified. Examples of quantitative data are GPA, service hours completed, or semesters enrolled full-time. The term *qualitative data* is used to describe data that approximates or describes but does not apply numeric measurement to define the characteristics or properties of a thing or a phenomenon. Examples of qualitative data are gender identity, college major, and hometown.

This distinction seems relatively straightforward. However, keep in mind that data is malleable. Sometimes a hometown is just a hometown. For example, someone identifies Boone, NC as their hometown and someone else identifies Greenville, NC as their hometown. These data are clearly qualitative. In addition, the data are nominal because there is no single right way to evaluate, rank, or order a list of hometowns.

Sometimes we take data like hometown and impose a schema on it…remote, rural, suburban, urban. There are now potential ways to think about ordering the categories…maybe in terms of population density. Although you can say more people live in a suburban square mile than in a rural square mile, based on these categories you can't really define what the exact difference might be between a rural and suburban location. We call data that can be classified in terms of a natural order ordinal data.

Sometimes we go a step further and impose a clearly quantifiable or measureable schema on our categorical data… in our current example we might classify towns by population size (e.g., 1 – 19,999, 20,000 – 29,999, 30,000 – 49,999, etc.). In this case these data are interval because the categories represent specific, equivalent measures.

This idea is important because the type of data you have has direct bearing on how you analyze and interpret that data. It doesn't make sense to calculate the average of a variable like college major. To draw meaning from it, you have to think about the nature of your data. What kind of questions can your data answer? What makes sense? Hopefully you have worked through this on the front end of the project, but it doesn't hurt to reconsider these questions once you have collected your data.

Interpreting Quantitative Data

In working with quantitative data in the context of ASSESSMENT, the focus is generally on describing patterns.

First, you may want to describe the characteristics or attributes of a whole group in this case we *aggregate* data (e.g., data for all UNC Students is analyzed together). At other times, you might want to know something specific about a particular group or groups. In this case you disaggregate your data (Analyze data for First Years, Sophomores, Juniors, and Seniors separately). You would use disaggregated data when we want to make group comparisons, or show change in a single group over time.

Next, when describing an individual variable (e.g., Campus Employment) you might choose to report things like:

- Frequency & Percentages – How many
- Range of Responses – Highest and lowest response
- Mean - Typical value or the average response
- Mode - Most frequently reported response
- Skew – Whether responses cluster around a particular point or side of the scale
- Criteria – Percent of students who meet or exceed an established criterion.

There are other times when you might want to describe the relationship between variables (e.g., Work on Campus & Retention). In these instances different statistics and statistical techniques are used (i.e., Correlation, Regression, Path Analysis). There are a number of resources available if these terms are new to you and you are interested in learning more. The full text of Beginning Statistics (Diamond & Jeffries, 2001) is available online via the UNC Library. The University of California at Berkeley offers an excellent free Introduction to Statistics course online (http://www.openculture.com/introduction_to_statistics_a_free_course). There is also a very good series of statistics courses offered through the School of Education on the UNC campus. The Odum Institute for the Social Sciences (also here on campus) offers statistical consulting as do members of the Assessment Council and the Coordinator for Assessment and Strategic Planning for Student Affairs.

Interpreting Qualitative Data

Qualitative data includes things like transcripts of focus groups or interviews, video or audio recordings, field or observation notes, journals, correspondence, websites, photographs, comments from an open-ended survey question, and other types of documents and/or artifacts. Like quantitative data analysis, goals of qualitative analysis include summarizing data, identifying patterns, and making meaning of the information to inform decision making and practice. Although the practice of interpreting qualitative data may vary somewhat in regards to the nature and purpose of an assessment project, *The 4 C's* (Conceptualizing, Coding, Categorizing, and Concluding) provide a simple means of describing and/or organizing the process.

Conceptualizing - Interpreting qualitative data begins with preparation and forethought. Although labor intensive, the first recommended step in working with qualitative data is to

produce a written transcript when appropriate (e.g., focus groups, interviews, video or audio recordings). Before analyzing any transcript, document, or artifact, the person or persons conducting analyses should establish how analysis will proceed. Will data analysis be grounded by a set of a priori or pre-established themes? Take the case of Orientation Leader training that heavily emphasized the development of a particular set of problem-solving strategies. In reading through OL incident reports (qualitative data), the person(s) analyzing the data might be looking for instances in which these particular strategies were enacted. In contrast to analysis in which you already know what you are looking for, there are times when open-ended exploration of a program, environment, or phenomena, etc. is justified. In these instances qualitative data analysis reflects an inductive process. Themes emerge from direct examination of the data. For example, you want to know how to structure a new program to maximize student engagement. In this case, it would be appropriate to host a series of focus groups, create transcripts or lists of student comments from these focus groups, and draw directly on participants' words to identify themes in the data that signify what is most meaningful to students. Once data have been prepared for analysis and the basic decision has been made regarding how patterns/themes in the data are to be identified (a priori, emergent) the process of coding can begin.

Coding – Saldaña (2009) defines a code as "a word or short phrase that symbolically assigns a summative, salient, essence-capturing, and or evocative attribute for a portion of language-based or visual data. (p. 3) In other words a code assigns a specified meaning to a segment of textual or visual data. As such the process of coding refers to the act of working through your data and marking appropriate segments of the data with symbols, descriptive words, or labels.

For example take the following passage:

Training gave me an opportunity to establish myself as part of a group. I was no longer on my own. I was a part of something bigger. It helped me understand that I was part of a network and a part of the University too.

Whereas several sentences make up the passage, each communicates the idea of feeling connected or having a sense of belonging. As such "connection" or "belonging" would represent an effective code to describe the theme of this passage.

The act of coding data represents an important step in analysis as coding establishes a framework for linking data to larger ideas. Very briefly, the process of coding involves transcribing and reading through data (textual or visual). When utilizing an emergent approach, the first read through should be used to get an overall sense of the data and to identify major themes/categories in the data that would represent meaningful codes. It is with the second reading that we read with the intent of identifying and assigning codes to meaningful segments of text. This process continues until all data have been read, reread, and meaningful segments coded.

Categorizing – Once initial coding has been completed and data are linked by organizing ideas or themes, it becomes possible to examine the patterns and or connections represented by the codes. The code itself implies meaning, but the act of coding makes deeper analysis possible. For example, coded data can be used to identify and understand:

- similarities or differences across actions, events, or phenomena, etc.
- frequency which can imply the importance of an action, event, or phenomena, etc.
- sequences or the order that things happen
- relationships between actions, events, or phenomena, etc.

In reality, the process of categorization involves a continued effort to summarize or reduce data through the process of refining your system of codes. You might identify codes that overlap to represent a larger theme or idea. Or you might notice ways in which codes can be organized hierarchically or instances in which one code usually occurs simultaneously with another. Groupings or overlap, points of conflict or tension, and examples of outliers are all things that might be of interest at this point in analysis. Categorization can also involve the "quantifying" of qualitative data. For instance, you might want to know how many times a particular code appeared in data provided by male participants versus how many times the code appeared in data provided by female participants. Generally speaking, this phase of analysis involves looking for broader patterns and or relationships in the data.

Concluding – Drawing conclusions from the analysis of qualitative data includes highlighting significant findings and discussing the possible meaning of findings in regards to future decision-making and practice. The following questions can help shape this process.

- What patterns, themes, or trends emerged?
- Were there notable group similarities or differences?
- Which themes occurred most or least frequently?
- Were there any noticeable relationships between themes?
- Were there outliers or instances that did not follow the expected pattern? Why?
- What stories emerge from the data?
- What questions remain? How will you answer them?
- How will findings be used to assist decision making or improve practice?