Texas Christian University
Student Affairs Assessment Guide

Chapter 5: Analyzing Assessment Results

It is important that student affairs professionals have the ability to articulate, interpret, and apply results of assessment reports and studies, including professional literature (NASPA). As Henning & Roberts (2016) assert, "Statistics are powerful tools in understanding students and their needs and improving our work. Complicated, complex statistics are fortunately not necessary for student affairs assessment…However, a basic understanding of statistics is useful in articulating why a statistical method was used or not used (p. 148)."

Statistics can be either descriptive or inferential. *Descriptive* statistics typically are used to describe a population. *Inferential* statistics are used to predict or draw conclusions about a population.

Levels of Data

In descending order of precision, the four different levels of measurement are:

- Nominal--Latin for name only (Republican, Democrat, Green, Libertarian); Categorical
- Ordinal--Think ordered levels or ranks (small--8oz, medium--12oz, large--32oz)
- Interval--Equal intervals among levels (1 dollar to 2 dollars is the same interval as 88 dollars to 89 dollars)
- Ratio--Let the "o" in ratio remind you of a zero in the scale (Day 0, day 1, day 2, day 3, ...)

The first level of measurement is **nominal level of measurement**. In this level of measurement, the numbers in the variable are used only to classify the data. In this level of measurement, words, letters, and alpha-numeric symbols can be used. Suppose there are data about people belonging to three different gender categories. In this case, the person belonging to the female gender could be classified as F, the person belonging to the male gender could be classified as M, and transgendered classified as T. This type of assigning classification is nominal level of measurement.

The second level of measurement is the **ordinal level of measurement**. This level of measurement depicts some ordered relationship among the variable's observations. Suppose a student scores the highest grade of 100 in the class. In this case, the student would be assigned the first rank. Then, another classmate scores the second highest grade of a 92; that student would be assigned the second rank. A third student scores an 81 and that student would be assigned the third rank, and so on. The ordinal level of measurement indicates an ordering of the measurements.

The third level of measurement is the **interval level of measurement**. The interval level of measurement not only classifies and orders the measurements, but it also specifies that the distances between each interval on the scale are equivalent along the scale from low interval to high interval. For example, an interval level of measurement could be the measurement of anxiety in a student between the score of 10 and 11, this interval is the same as that of a student who scores between 40 and 41. A popular example of this level of measurement is temperature in centigrade, where, for example, the distance between 94^0C and 96^0C is the same as the distance between 100^0C and 102^0C.

The fourth level of measurement is the **ratio level of measurement**. In this level of measurement, the observations, in addition to having equal intervals, can have a value of zero as well. The zero in the scale

makes this type of measurement unlike the other types of measurement, although the properties are similar to that of the interval level of measurement. In the ratio level of measurement, the divisions between the points on the scale have an equivalent distance between them.

It should be noted that among these levels of measurement, the nominal level is simply used to classify data, whereas the levels of measurement described by the interval level and the ratio level are much more exact.

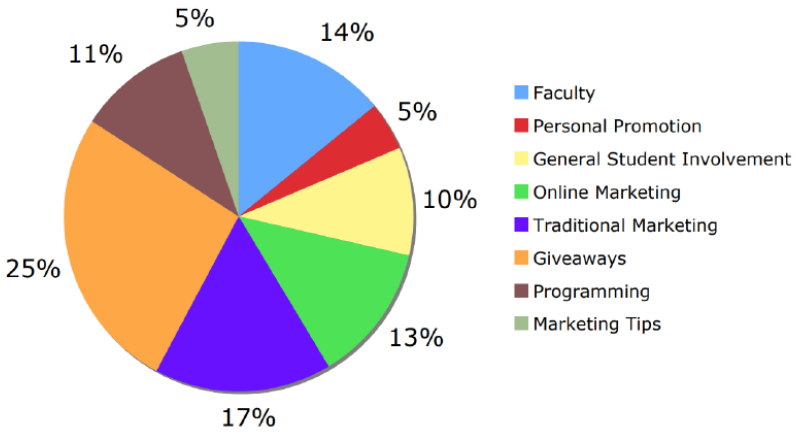
Descriptive Statistics

Descriptive statistics are used to describe the basic features of the data in a study. They provide simple summaries about the sample and the measures. Together with simple graphics analysis, they form the basis of virtually every quantitative analysis of data. Descriptive statistics do not, however, allow for conclusions to be made beyond the data that has been analyzed or reach conclusions regarding any hypotheses that might have been made. They are simply a way to describe our data.

Measures of central tendency and measures of dispersion are the two types of descriptive statistics. The mean, median, and mode are three types of measures of central tendency. The range, variance, and standard deviation are three types of measures of dispersion.

Frequency tables and graphs are indicative of descriptive techniques. The following illustration is a prime example.

What is the most effective way to get students to attend an event on campus?



Frequency Table	
Faculty Involvement	110
Personal Promotion	35
General Student Involvement	77
Online Marketing	98
Traditional Marketing	128
Giveaways	204
Programming	82

Measures of central tendency:

- Mean—average of a set of numbers
- Median—middle number in a list of numbers
- Mode—most frequently occurring number in a set of numbers

Measures of dispersion:

- Range—distance between the smallest and the largest number
- Variance—statistical average of the amount of dispersion in a set of numbers
- Mode—measure of the extent to which numbers are distributed around the mean (square root of variance)

Inferential Statistics

Inferential statistics are used to infer from the sample data what the population might think or do. Inferential statistics are used to reach conclusions that extend beyond the immediate data alone. Inferential statistics are used to make judgments of the probability that an observed difference between groups is a dependable one or one that might have happened by chance. Generalizations about the larger population may be drawn from a smaller sample by using inferential statistics.

Basic Concepts of Quantitative Analysis

In being able to "speak the language," there are a few basic concepts it is important to understand:

Confidence Interval: provides a range within which the population mean is likely to occur. Aka: Margin of error. Rule of thumb, 4% is standard confidence interval. For example, someone says, "75% of students loved it, +/- 4%." This means the confidence interval is + or – 4 points, so the "real" percent of students who loved it is somewhere between 71 and 79%.

Effect Size: helps determine the practical significance of the obtained results. It is a statistical concept that measures the strength of the relationship between two variables on a numeric scale. An effect size of 0.3 is considered small, 0.5 moderate, and 0.7 large.

Alpha Level: the odds that the observed result is due to chance. A .05 alpha means that 5 times out of 100 the obtained results are due to chance rather than the 'treatment.' Aka: Significance level. (%)

Power: the odds that the observed results are, in fact, due to the 'treatment.' (1-%)

Parametric Tests: assume that the sample is normally distributed, randomly selected, and representative of the population. Recommended when the data is interval or ratio level. Examples include t-tests, ANOVA, and Regression.

Nonparametric tests: Make no assumptions are made about the distribution of the data. Recommended when the data is nominal or ordinal level. Examples include Chi-square, Wilcoxen Signed Rank, and Kruskal-Wallis.

Reliability: level of confidence that a tool used to gather data is consistent over time.

Internal Validity: level of confidence that a tool used to gather data measures what it says it measures.

External Validity: level of confidence that obtained results can be generalized to the population or other situations.

Multicultural Validity: level of confidence that the tool used to gather data is unbiased across dimensions of cultural diversity.

Generalizability: extent to which obtained results can be applied to the population the sample represents.

Qualitative Data Analysis

Qualitative data analysis seeks to tell the story of a particular group's experiences in their own words, and is therefore focused on narrative rather than on numbers. Unlike quantitative research, in which researchers state specific hypotheses and then collect data to empirically test them, most qualitative research employs an inductive approach in which the researcher first collects data and then attempts to derive explanations from those data. As such, qualitative research tends to be more exploratory in nature, seeking to provide insight into how individuals (or organizations, groups, etc.) understand aspects of their worlds.

Often critiqued as being "soft," and open to multiple interpretations (in contrast to the hard numbers associated with quantitative research), qualitative analysis is frequently misjudged. Well-designed qualitative research is highly systematic, however, requiring that researchers carefully record both their observations and their experiences in collecting the data. A related concern is that qualitative research tends to confirm researchers' own understanding of phenomenon. Good qualitative researchers aim for "saturation" of data, however, which involves collecting and analyzing data in an iterative process until no new information arises. Systematically collected, saturated qualitative data provide a degree of insight into complex phenomenon that differs dramatically from what is provided through quantitative analysis, and both approaches should be valued for the unique contributions they make in research.

Throughout the course of qualitative analysis, the analyst should be asking and reasking the following questions:

- What patterns and common themes emerge in responses dealing with specific items? How do these patterns (or lack thereof) help to illuminate the broader study question(s)?
- Are there any deviations from these patterns? If yes, are there any factors that might explain these atypical responses?
- What interesting stories emerge from the responses? How can these stories help to illuminate the broader study question(s)?
- Do any of these patterns or findings suggest that additional data may need to be collected? Do any of the study questions need to be revised?
- Do the patterns that emerge corroborate the findings of any corresponding qualitative analyses that have been conducted? If not, what might explain these discrepancies?

Broadly speaking, qualitative data analysis can be broken down into three broad tasks: data reduction, data display, and conclusion drawing with verification (Miles & Huberman, 1994). After data has been collected through interviews, focus groups, observation, or social media, the process of

Data reduction refers to the process of selecting, focusing, simplifying, abstracting, and transforming the data that appear in written up field notes or transcriptions. Not only do the data need to be condensed for the sake of manageability, they also have to be transformed so they can be made intelligible in terms

of the issues being addressed. Data reduction should be guided primarily by the need to address the salient evaluation question(s). Data reduction often forces choices about which aspects of the assembled data should be emphasized, minimized, or set aside completely for the purposes of the project at hand.

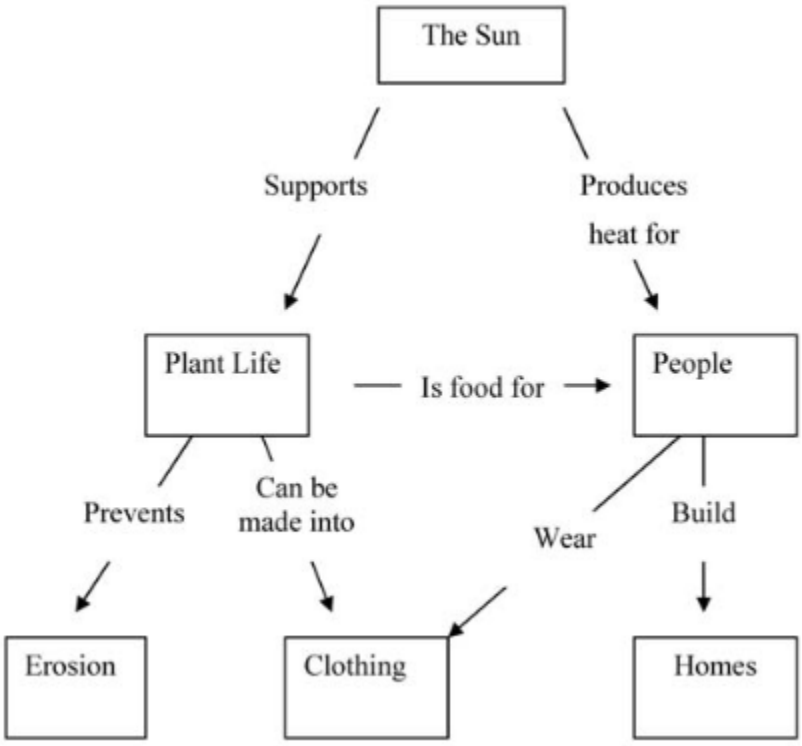
Most people refer to data reduction as 'coding.' Codes serve as a way to label, compile and organize the data. Codes also provide a way to summarize and synthesize what is happening in the data. In linking data collection and interpreting the data, coding becomes the basis for developing the analysis. It is generally understood, then, that "coding is analysis." Schuh (2016) provides a brief seven step process for coding data:

Data display goes a step beyond data reduction to provide an organized, compressed assembly of information that permits conclusion drawing... A display can be an extended piece of text or a diagram, chart, or matrix that provides a new way of arranging and thinking about the more textually embedded data. Data displays, whether in word or diagrammatic form, allow the analyst to extrapolate from the data enough to begin to discern systematic patterns and interrelationships. At the display stage, additional, higher order categories or themes may emerge from the data that go beyond those first discovered during the initial process of data reduction. From the perspective of program evaluation, data display can be extremely helpful in identifying why a system (e.g., a given program or project) is or is not working well and what might be done to change it. One technique for displaying narrative data is to develop a series of flow charts that map out any critical paths, decision points, and supporting evidence that emerge from the data. Using a data matrix may help when comparing groups of respondents. A data matrix is a two dimensional tabular representation of data in which data fields can be organized by rows and columns. The point of intersection between a row and column is a cell. The following table is an example of a data matrix from a study investigating how knowledge is shared across campus.

Exhibit 10.
Data matrix for Campus A: What was done to share knowledge

Respondent group	(a) Activities named	(b) Which most effective	(c) Why
Participants	• Structured seminars • E-mail • Informal interchanges • Lunchtime meetings	• Structured seminars • E-mail	• Concise way of communicating a lot of information
Nonparticipants	• Structured seminars • Informal interchanges • Lunchtime meetings	• Informal interchanges • Structured seminars	• Easier to assimilate information in less formal settings • Smaller bits of information at a time
Department chair	• Structured seminars • Lunch time meetings	• Structured seminars	• Highest attendance by nonparticipants • Most comments (positive) to chair

The software package NVivo, developed to facilitate qualitative analysis, provides the ability to construct concept maps and mind maps of identified themes, categories of respondents, and cases. In general terms, concept mapping is a technique that can demonstrate how people visualize relationships between various concepts (Lanzing, 1996). The illustration below provides an example of a simple concept map.



The Framework Method (Ritchie & Spencer, 2003) is becoming an increasingly popular approach to the management and analysis of qualitative data. The Framework Method sits within a broad family of analysis methods often termed thematic analysis or qualitative content analysis. These approaches identify commonalities and differences in qualitative data, before focusing on relationships between different parts of the data, thereby seeking to draw descriptive and/or explanatory conclusions clustered around themes.

Stage 1: Transcription

A good quality audio recording and, ideally, a *verbatim* (word for word) transcription of the interview is needed. Transcripts should have large margins and adequate line spacing for later coding and making notes. The process of transcription is a good opportunity to become immersed in the data and is to be strongly encouraged. However, in some situations, the decision may be made that it is a better use of resources to outsource this task to a professional transcriber.

Stage 2: Familiarization with the interview

Becoming familiar with the whole interview using the audio recording and/or transcript and any contextual or reflective notes that were recorded by the interviewer is a vital stage in interpretation. It can also be helpful to re-listen to all or parts of the audio recording. One margin can be used to record any analytical notes, thoughts or impressions.

Stage 3: Coding
After familiarization, the researcher carefully reads the transcript line by line, applying a paraphrase or label (a 'code') that describes what they have interpreted in the passage as important. In more inductive studies, at this stage 'open coding' takes place, i.e. coding anything that might be relevant from as many different perspectives as possible. Codes could refer to substantive things (e.g. particular behaviors, incidents or structures), values (e.g. those that inform or underpin certain statements, emotions (e.g. sorrow, frustration, love) and more impressionistic/methodological elements (e.g. interviewee found something difficult to explain, interviewee became emotional, interviewer felt uncomfortable). In purely deductive studies, the codes may have been pre-defined (e.g. by an existing theory, or specific areas of interest to the project) so this stage may not be strictly necessary and the researcher could just move straight onto indexing, although it is generally helpful even if you are taking a broadly deductive approach to do some open coding on at least a few of the transcripts to ensure important aspects of the data are not missed. Coding aims to classify all of the data so that it can be compared systematically with other parts of the data set. As well as getting a holistic impression of what was said, coding line-by-line can often alert the researcher to consider that which may ordinarily remain invisible because it is not clearly expressed or does not 'fit' with the rest of the account. In this way the developing analysis is challenged; to reconcile and explain anomalies in the data can make the analysis stronger. Coding can also be done digitally using *NVivo*, which is a useful way to keep track automatically of new codes. However, some researchers prefer to do the early stages of coding with a paper and pen, and only start to use *NVivo* once they reach Stage 5.

Stage 4: Developing a working analytical framework
After coding the first few transcripts, researchers should meet to compare the labels they have applied and agree on a set of codes to apply to all subsequent transcripts. Codes can be grouped together into categories (using a tree diagram if helpful), which are then clearly defined. This forms a working analytical framework. It is likely that several iterations of the analytical framework will be required before no additional codes emerge. It is always worth having an 'other' code under each category to avoid ignoring data that does not fit; the analytical framework is never 'final' until the last transcript has been coded.

Stage 5: Applying the analytical framework
The working analytical framework is then applied by indexing subsequent transcripts using the existing categories and codes. Each code is usually assigned a number or abbreviation for easy identification (and so the full names of the codes do not have to be written out each time) and written directly onto the transcripts. *NVivo* is particularly useful at this stage because it can speed up the process and ensures that, at later stages, data is easily retrievable. It is worth noting that unlike software for statistical analyses, which actually carries out the calculations with the correct instruction, putting the data into a qualitative analysis software package does not analyze the data; it is simply an effective way of storing and organizing the data so that they are accessible for the analysis process.

Stage 6: Charting data into the framework matrix
Qualitative data are voluminous (an hour of interview can generate 15–30 pages of text) and being able to manage and summarize (reduce) data is a vital aspect of the analysis process. A spreadsheet is used to generate a matrix and the data are 'charted' into the matrix. Charting involves summarizing the data by category from each transcript. Good charting requires an ability to strike a balance between reducing the data on the one hand and retaining the original meanings and 'feel' of the interviewees' words on the other. The chart should include references to interesting or illustrative quotations. These can be

tagged automatically if you are using *NVivo* to manage your data. *NVivo* provides mechanisms to facilitate this stage of the process.

Stage 7: Interpreting the data
It is useful throughout the research to have a separate notebook or computer file to write down impressions, ideas and early interpretations of the data. Gradually, characteristics of and differences between the data are identified, perhaps generating typologies, interrogating theoretical concepts (either prior concepts or ones emerging from the data) or mapping connections between categories to explore relationships and/or causality. If the data are rich enough, the findings generated through this process can go beyond description of particular cases to explanation of, for example, reasons for the emergence of a phenomena, predicting how an organization or other social entity is likely to instigate or respond to a situation, or identifying areas that are not functioning well within an organization or system. It is worth noting that this stage often takes longer than anticipated and that any project plan should ensure that sufficient time is allocated to meetings and individual researcher time to conduct interpretation and writing up of findings

Research Literacy
Research literacy entails knowing:

- Basic principles of qualitative and quantitative research
- How research articles are put together
- How to read research articles at increasingly complex levels
- How to evaluate the quality of the research

A research article typically follows this outline:

- Title
- Names and Institutional homes of researchers
- Abstract
- Introduction
- Methods/Procedures
- Results/Findings
- Discussion/Conclusion
- References

A good abstract will include the purpose of the study, a statement of the problem, description of participants, depiction of the study design, explanation of data analysis procedures, summary of the findings, and presentation of conclusions that were drawn from the results.

When thinking about the purpose of research, there are four main types of research:

- Exploration – where things are new or poorly understood;
- Extension – studies built on other studies;
- Expansion – extends the work into new or more complex areas; and,
- Correction – earlier research is wrong, and this sets the record straight.

The rationale of a research project should answer the “so what?” question. Five primary rationales have been identified. These are:

- Crisis – something needs to change to avoid problems;
- Importance – elevates the topic to new level of practical significance;
- Gap Filling – topic already considered important, this project just fills in the gaps of knowledge;

- Depth – delve deeper into existing knowledge (often qualitative); and
- Commitment – explores area of inequity (often qualitative participatory action research).

The methods section of a research article should answer the "Who," "What," "How," "When," and "Where" of the project.

- Who – sampling and subjects;
- What – data gathering and instrumentation;
- How – research design and analysis procedures;
- When – longitudinal, pre/post;
- Where – setting or location;

The layout and composition of the findings section of a research article will typically depend on whether the project was quantitative or qualitative in nature. Usually results are presented by way of narrative combined with tables, graphs, and charts. Findings should be directly tied to the research questions.

While the "Findings" section presents the results that were found, the "Discussion" section allows for interpretation of the results. This section should help explain what the results mean in the broader context. This section may also include suggestions for future projects.

Hopefully, the information included in each of the chapters in this guide will contribute to gains in research literacy.

ACKNOWLEDGEMENTS

A huge “thank you” to CampusLabs for providing much of the information in this guide. The Accreditation course provided a great framework and many of the resources in this text.