Section 7: Data Analysis

Table of Contents

Section 7: Data Analysis

Intended learning outcomes

This section is designed to acquaint the implementation team with data analysis. The intended learning outcomes follow.

Upon completion of this section, the implementation team will be able to:

1. Describe advantages and limitations of qualitative, quantitative, and mixed methods analysis.
2. Understand ways that these strategies have been used and can be used to study specific STD-related topics.
3. Perform streamlined and efficient data analysis.

Chapter 1 introduces the topic of data analysis.

Section 7, Chapter 1: Introduction to data analysis for the Rapid Ethnographic Assessment

7.1.1 Introduction to data analysis for the REA

Nothing in research is as exciting as discovering what has been found as a result of all the data collection. When patterns begin to emerge, all of the effort the implementation team has put into the study becomes worthwhile. At times the team may begin to believe that nothing noteworthy will be discovered in the study, or the results will be too confusing to interpret. It is easy to begin believing this while collecting data because just about the time the team begins believing the data are suggesting one pattern, a series of interviews will put a damper on that.

But at Jill Florence Lackey & Associates, we have rarely collected data where some relevant patterns failed to emerge. We have often had unexpected results, but we have almost never had findings that could not be interpreted in ways that increased knowledge of a particular phenomena.

Data analysis is the bridge to interpreting findings. Analysis ultimately involves reducing the quantity of data to manageable chunks, but in ways that best summarize and explain findings. Creswell (2003) provides the following definition.

> *The process of data analysis involves making sense out of the text and image data. It involves preparing the data for analysis, conducting different analyses, moving deeper and deeper into understanding the data, and making an interpretation of the larger meaning of the data* (p. 190).

This definition implies several steps, which include (a) preparing data for analysis, (b) use of specific analytic techniques, (c) presenting the data in ways that move deeper into understanding the findings, and (d) interpreting the data. Throughout this section, these steps will be discussed, whether dealing with quantitative or qualitative data.

However, the reader should be aware that data analysis is a very complex topic. This section will focus on the more rudimentary forms of analysis. These forms can be easily learned and executed by team members who will be included in the analysis of data but have little or no experience with the process, such as with community team members in studies using a community-based participatory research (CBPR) approach. Some complex forms will be mentioned or introduced, but the explanations will only include very streamlined versions of these more complex types of analysis—usually those strategies that would be most appropriate for the REA. The instructions will emphasize those forms of data analysis used in ethnographic research and those forms that can be accomplished rapidly.

Data analysis is an enjoyable exercise. Do have fun with it.

Section 7, Chapter 1: Introduction to data analysis for the Rapid Ethnographic Assessment

7.1.2 Resources

Chapter references

Creswell, J.W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches* (2nd ed.). Thousand Oaks, CA: Sage.

Section 7, Chapter 2: Qualitative data analysis

7.2.1 Intended learning outcomes

The intended learning outcomes of this chapter on qualitative data analysis follow.

At the end of this chapter the implementation team will be able to:

1. Use three forms of qualitative data analysis;
2. Perform four steps in the forms of qualitative data analysis, including
   a. Preparing data for analysis
   b. Performing analysis
   c. Displaying results of analysis
   d. Interpreting findings;
3. Follow procedures in qualitative analysis in a systematic way;
4. Perform quality checks on the analysis.

Section 7, Chapter 2: Qualitative data analysis

7.2.2 Introduction

Tashakkori and Teddlie (1998) define qualitative data analysis in the following way.

> *The essence of qualitative data analysis of any type is the development of a typology of categories or themes that summarize a mass of narrative data* (p. 119).

Thus, the emphasis in qualitative data analysis is reducing a mass of information mainly through categorization. This chapter will focus on streamlined ways to categorize. In nearly every example of qualitative research, findings can be categorized in ways that can be presented in narration form (such as through stories that emerge from observation notes or interview/focus group excerpts) or in tabular or graphic forms (such as outlines, matrices, taxonomies, networks, or flow charts).

There are many styles and theories associated with qualitative data analysis, and the topic in itself is highly complex. However, this curriculum is not for professional researchers or for students entering a field of research. It is designed to teach relative novices how to conduct "ethnography" "rapidly" in order to assess issues related to STD prevention/treatment and STD introduction/transmission. This chapter will thus focus on qualitative data analysis designed for ethnography. In the majority of cases (although there are clear exceptions), ethnographers take an inductive approach to collecting and analyzing data[1]. Categories of meaning often emerge as the study is being conducted, as opposed to deductive strategies where the categories are defined before the study begins. In the REA, the focus has been on asking research questions rather than testing hypotheses (although experimental designs are deductive at least in their final stages).

But even asking questions can pre-categorize some of the information that is gathered. Participants do respond to specific questions in interviews. Some of the observation protocols the team has developed have places on the pre-printed guides where the observer watches for specific topics. But generally, the assessment has been designed to be inductive in approach because the implementation team needs information from the "insider's perspective," as stated in the definition of RAP (Beebe, 2001) and our REA definition (see the first section, The Problem and Target Population).

Not discussed in this chapter are excellent forms of qualitative data analysis associated with other research traditions, such as grounded theory, phenomenological studies, and narrative research ("restorying participants' stories). Those interested in these forms can refer to the list of additional resources listed at the end of the chapter.

***Terms*.** Specialists on qualitative data use a variety of terms associated with the "categories" that are identified in data analysis. Labels for these categories include "codes," "indexes," "variables," "terms," "items," "topics," and "domains." Some of these terms have subtle differences in application, but because a purpose of the curriculum is to produce reader-friendly and streamlined instruction in the REA, we will not get into the more complicated discussions that require distinctions. Rather, we will follow a certain pattern in use of these terms, and a pattern that we believe specialists in qualitative data analysis most commonly apply.

When we are discussing the process of categorizing the data, we will use the term "coding," which involves "selecting, focusing, simplifying, abstracting, and transforming the data" (Beebe, 2001, p. 66). The labels for the categories are "codes." But once the categorizing process is complete, we will refer to each category of data as a "variable." According to Pelto & Pelto (1987), a "variable" is "any of the elements, concepts, or categories in theoretical propositions that are thought to vary in degree or in kind" (p. 142). For the REA purposes, this basically means that the categories of information that the implementation team identifies will end up being part of some basic scheme/system/theory in which the findings are summarized and interpreted. And keep the word "varies" in the back of the mind while reviewing this chapter.

Another important term is "taxonomy." A taxonomy is a set of categories organized on the basis of a single relationship, such as "types" of things or "reasons" for doing something (Spradley, 1980, p. 112). Related to taxonomies is a "matrix" (plural, "matrices"). A matrix is the "crossing" of two lists, set up as rows and columns in order to show comparisons (Miles & Huberman, 1994, p. 93).

Advantages and limitations of qualitative data analysis

Advantages. Many advantages of qualitative data analysis are advantages of qualitative inquiry generally. Qualitative data analysis (particularly when codes have not been predetermined) can help identify completely unexpected themes and patterns. This contrasts with quantitative data analysis, where the researcher is summarizing information from closed-ended questions. The unexpected findings gleaned through qualitative analysis could end up being the impetus that checks an STD outbreak in some area.

Second, qualitative data analysis helps researchers uncover the meanings that participants attribute to phenomena. Meanings are extremely important in STD-affected communities. Some questions might be: What are folk beliefs associated with STD-preventive measures? How do people size up potential partners? What criteria do they use in suspecting that potential partners may be infected with STDs? What criteria do they use to determine whether potential partners are not infected with STDs?

Presentation and display of qualitative findings can also be accomplished in many interesting ways. For example, through use of excerpts from interviews, audiences can hear insider perspectives in the participants' own words. Use of quotes is particularly important in demonstrating assumptions people may have about STDs that might not be identified through survey research.

In addition, many audiences may not care to hear all results in numbers. Some people think statistics are boring, untrustworthy, or just plain dehumanizing, Those presenting qualitative findings can tell stories or provide case study examples that would be more appealing to these kinds of audiences.

Researchers can also analyze qualitative data by using diagrams, charts, and taxonomies that can help audiences "see" the wide-angle view of the phenomena under study. Taxonomies could categorize types of STD services; flow charts could demonstrate processes in STD introduction; maps could show how and where an STD is passing through a population.

***Limitations*.** Possibly the greatest limitation of qualitative data analysis is the way the "typologies" that the analysis creates (per the definition of Tashakkori and Teddlie) can stereotype a full population. The diagrams tend to illustrate processes, meanings, and relationships in ways that might suggest that all members of the population are "common denominator" people. In quantitative analysis one can show, say, percentages of people that follow the pattern.

Another limitation of note is the amount of subjectivity that can be involved in qualitative analysis. While there are ways that the subjectivity can be reduced, it is ultimately the researcher that makes decisions about the way the information is categorized and presented, or how the data are interpreted. In quantitative analysis the researchers also make some decisions (particularly about interpretation), but the opportunities to make subjective decisions are fewer.

There are also limitations associated with display and presentation. It can take a long time to explain both the way analysis is done and the actual findings to audiences. For example, if convincing an audience of a particular interpretation depends on reciting stories from observation and interview excerpts, the researcher must provide a substantial number of examples to be plausible.

A word about computer programs

There are a variety of computer programs available to analyze qualitative data. A table on these programs is provided in the appendix. The table gives information on strengths and weaknesses of the programs.

The implementation team may have access to these programs or may wish to purchase one. However, we are not recommending this for the Rapid Ethnographic Assessment. Unless a member of the team is already familiar with their use, the learning time may be prohibitive for a study of this duration. In addition, Beebe (2001) states that the paper and pencil form of analysis is appropriate for rapid assessment, and Bernard (2006) argues that the software programs only begin to be advantageous when thousands of pages of (typed) qualitative data have been generated (which we clearly do not anticipate in the REA).

Wherever computer programs are used in data collection and analysis (even in simple word processing), the implementation team *must* back up all files on diskettes or CDs, as well as print out hard copies. Ethnographers share horror stories about failure of hard drives and software and loss of laptops in the field.

Levels of qualitative analysis

The levels of difficulty presented throughout this curriculum are designed to alert the implementation team to the time that may be involved in learning and implementing procedures early in the REA planning process. However, the actual procedures selected *must be selected because they would best answer the questions that the REA is asking or the information being gathered, not the level of difficulty*.

We will address three levels of qualitative analysis in this chapter. The first is simple topical coding.

Simple topical coding

This first level of qualitative analysis described below would be appropriate for times when the implementation team has a strong quantitative section in the study and the qualitative data are designed to be mainly supportive or complementary to the quantitative. Topical coding is what the label implies—categorizing data by topic (Bernard, 2006, pp. 399-404). Topics can be an endless variety of phenomena, such as types of places, attributes of people, reasons for doing things, strategies, assumptions, and kinds of attitudes or beliefs. Coding, as previously mentioned, is the process of breaking data into chunks, organizing and ordering data into categories, and adding labels to the categories.

***Preparing data for analysis*.** Coding should not begin when all the data are collected. Coding should be an ongoing process begun relatively early in the data collection stage. The interview and observation protocols that the implementation team has developed had places on them where the researchers could write down ideas, notes, hunches, and possible patterns. This is the beginning of the coding process.

The next stage is word-processing hand-written notes from observation and interviews/focus groups and transcribing video- and/or audiotapes. There may be times when the implementation team has so little time to complete the REA that they may forego word-processing the hand-written notes (tapes must always be transcribed if one is to analyze the data). On occasion, we at Jill Florence Lackey & Associates have simply not had the time to type hand-written notes. If this does occur, then coding will have to be done on the hand-written material. However, we do not recommend this for two reasons. First, more than one researcher will be involved in coding and handwriting is often hard to read—particularly when writing is done in a hurry. Second, there is hardly a better way to re-familiarize oneself with the data than to type the hand-written notes, and hardly a better time to familiarize oneself with interviews and observations conducted by other team members.

Assuming that team members will try to find the time to type material from hand-written sources (and of course from tapes), the team should allow for ample margin room for marginal codes to be inserted. One might even format the pages into columns with the widest of the columns being used for the data and one or two narrower columns being used for coding and remarks. This way codes can be typed onto the pages, if this makes the job easier. Typing should be double-spaced.

Performing the analysis: categorizing the data**.** At an agreed upon interval during data collection, members of the implementation team should begin coding their qualitative texts (use of any other visual material such as photographs should also be coded by topics). This interval should occur when a good representation of the data is collected (often when the researchers are about one-quarter of the way through). A sample of these data is then selected for the pre-coding process. More than one member of the implementation team should be involved in this stage. These researchers, working independently, should look for broad categories ("umbrellas") of information in the texts, which can later be subdivided. For example, if the study is supposed to assess the level of services available for STD treatment in a certain area, umbrella codes might be "services for viral infections" and "services for bacterial infections." Most specialists in data analysis (e.g., Miles & Huberman, 1994; Bernard, 2006) recommend beginning with as few umbrella codes as possible (five or six), as the list tends to expand over time—particularly subcategories. Once the researchers have gone through the data sample and identified their initial list of umbrella codes, they should come together and see how they agree on the codes. Where disagreement occurs, compromises can then be sought.

Now would be a good time to watch the video on Getting started with coding. The video demonstrates ways that a team can initially identify codes and how the team can reach agreement on the final categories.

The team now has its starting codes, which can then be subdivided. But first the team should label the codes. There are two common ways codes are labeled—by numbers or by letters. Ethnographers traditionally used numbering systems, where subcodes could be labeled by adding decimal points to the original number. However, many have switched to a lettering system where abbreviations of words can be used. Miles & Huberman argue that numbers are too hard to remember. They employ a system where abbreviated words can be used, followed by dashes and slashes for the subdivisions. For example, the above mentioned study of STD services might employ the umbrella codes of VIR for services relating to viral infections and BAC for services relating to bacterial infections, and then subdivide by the specific kind of infection, followed by additional subdivisions if necessary. One subdivided code might then be VIR—hiv/aids, if the researchers need to distinguish full-blown AIDS from HIV.

A note of caution about categorizing . . .

The members of the implementation team should be aware that the categories and subcategories they are choosing as meaningful are their own. This is seldom a contested issue when the categories are subdivisions of classification systems that are fairly universally accepted, such as distinguishing viral infections from bacterial infections, and distinguishing the specific subtypes of infections. However, this changes when the implementation team begins categorizing data that reflect the assumptions and beliefs of the target community. The researchers should always be aware that the development and identification of codes, themes, patterns, and taxonomies is always subjective at some level.

The “emergent research sequence” mentioned later in this chapter is much more complicated than simple topical coding, but helps the researcher, to some greater degree, identify the way the target community categorizes phenomena.

The codes the implementation team has selected are then written next to the qualitative text, and can be color-coded by use of markers, if the team wishes to visualize easily the start and end point of the segment. Because there are a limited number of colors in markers, the color-coding should probably only include the “umbrella” codes, not the subcodes. If the codes are being typed on the pages, the text can be shaded, if desired. See the following example from the hypothetical study on STD prevention and treatment services. The text is from a mock interview.

"I believe we need to look at what is really being done in terms of outreach	BAC--syph
to people with syphilis. We had an outreach specialist for that at the clinic years	
ago. But today the emphasis in this area has been strongly on HIV, and people have	VIR-hiv
lost track of the point that these other infections are still out there. It is truly	
my opinion that some of our public health specialists need some serious re-training on	BAC-syph
syphilis and, to some degree, on gonorrhea. I hear all the misinformation. But I don't	BAC-gon
see it happening."	

Patterns emerge where the researchers see the same kind of information repeated again and again on the specific coded topics. For example, if nearly every time the code of BAC-syph appears, the person interviewed defines some kind of service need, this would constitute a pattern.

In addition, other notes can be marked in the margins, and any reflection notes that appeared in the original hand-written notes should be repeated in the margins. What we also tend to do at Jill Florence Lackey & Associates is highlight certain segments of the text in boldface if we think we may later want to use the segments as typical (and well-articulated) examples in our later write-ups or presentations.

As the subcategories of codes grow, the implementation team will need to create a "code book." The team will want to agree on some organizing scheme to keep track of the code labels and specific definitions. Some researchers using numerical codes order the codes in ascending numbers and others using letter codes order the codes alphabetically. We suggest that the team begins by organizing the codes under the limited number of umbrella codes, and then reach consensus on how to order the subcategories under the larger umbrellas.

But how should one organize the texts the codes are representing? The implementation team can keep a log of the coded material in a variety of ways. Some researchers use their word processing programs to cut and paste the excerpts into files organized by code labels. Others cut photocopied excerpts with scissors and paste them onto 3x5 cards (or larger). The latter is particularly useful when the texts are still in the hand-written forms (but do have the originals photocopied). Minimally, the team should keep a list of the codes and under each code list the page numbers and documents where the excerpts appear (so the researcher can locate the exact text later), and make notes where some particular excerpt has a feature of interest. Keeping these logs will also later help the researchers count the number of times the codes appear in the texts overall, and in certain contexts in particular.

However these excerpts are archived, the team should include in the excerpts the following information about each chunk of data:

- The code label,
- The site where the data were collected,
- The page number from the original documents,
- The person speaking (where applicable),
- The focus group (where applicable)

Depending on the kind of information the implementation needs, other information might be included, such as date, event, and researcher's name.

***Refining the analysis: Display/presentation of data*.** Once the coding process is complete, we will refer to the data categories as "variables." An easy way of deciding which variables the team wishes to work with in displaying and later interpreting the data is simply by counting the number of times text excerpts appear under the codes. This is an easy way to begin. This counting process is described by Miles & Huberman (1994).

In qualitative research, numbers tend to get ignored…However, a lot of counting goes on in the background when judgment of qualities are being made. When we are identifying a theme or a pattern we are isolating something that (a) happens a number of times and (b) consistently happens in a specific way. The "number of times" and "consistency" judgments are based on counting. When we make a generalization, we amass a swarm of particulars, and decide, almost unconsciously, which particulars are there more often, matter more than others, go together, and so on. When we say something is "important" or "significant" or "recurrent," we have come to that estimate in part by making counts, comparisons, and weights. . .

There are three good reasons to resort to numbers: to see rapidly what the team has in a large batch of data; to verify a hunch or hypothesis, and to keep oneself analytically honest, protecting against bias. (p. 253)

Thus a good place for the implementation team to start working with variables is to count the number of times they are coded in the qualitative texts. This does not mean that one needs to ignore the variables that come up less frequently, but the team should first work with the more numerous ones ("to keep yourself analytically honest" per the above citation). However, you do not want to reduce the analysis to a count of frequency of specific codes. Qualitative approaches allow depth of understanding rather than a mere count of the frequency of the phenomena in question.

One easy way to display the data is simply by citing the more numerous variables and their variations in a text form, and then using excerpts from interviews, life histories, focus groups, or observation notes. For example, suppose the implementation team was studying an STD-at-risk community through focus groups and asking participants to discuss STD prevention measures. Suppose that participants in each focus group spent considerable time debating whether or not they should get AIDS testing. If this was the case, it is likely that a code in textual analysis would be something like "reasons why get AIDS testing" and "reasons why not get AIDS testing," and then these codes would be subdivided into the categories of reasons given. A data summary on "reasons why not get AIDS testing" might then look something like the following example.

In our focus groups conducted at the Barns and Haley community centers, participants frequently debated the topic of AIDS testing. Most of those who said they would never get testing gave the following reasons: (1) fear of consequences or (2) certainty that they did not have the disease. Of those citing the "fear factor," the largest number simply said they were afraid of finding out they had the disease. See below.

"[Gp. 7, #2] If I do have it, I don't want to know."

"[Gp. 5, #8] If I have it and don't know it, I live on like a normal person. Why would I want to know I have it? I think denial can be a good thing [laughter by other participants]."

Others claimed they were afraid of being seen while getting the test.

"[Gp. 1, #5] I might be at that clinic and some street guy knows me, then tells the whole world I think I got AIDS."

"[Gp. 3, #9] Someone sees me in that waiting room and I'm outed."

Still others expressed concerns that confidential records would be made public.

"[Gp. 1, #10] I don't believe it when they say it is anonymous. Don't tell me the testing people don't say something when they see you later on the streets."

The arguments of those who fell under the category of "certainty they did not have the disease" were based more on their specific histories (some including misinformation) than on emotions.

"[Gp. 2, #5] It's pointless. I've been with the same man and no one else for 10 years. He sure doesn't have the [disease][2]. None of us do IV drugs."

"[Gp. 3, #3] I would know it if I was HIV positive. I would be having sores or something."

"[Gp. 3, #7] I have a wife. She's monogamous. I watch who I fool around with."

What has actually been developed above is a kind of taxonomy. The same information can be placed in an outline or tabular form. This gives the audience an interesting way of following the findings that can help them see the categories more clearly. In the taxonomy, a limited number of the actual quotes can still appear. The following two examples show the outline form and a table.

1. Reasons given why participants would not get AIDS testing

   A. Fear of consequence

      1. Fear of learning they were HIV positive

         "[Gp. 7, #2] If I do have it, I don't want to know."

      2. Fear of being seen getting test

         "[Gp. 3, #9] Someone sees me in that waiting room and I'm outed."

      3. Fear of records being made public

         "[Gp. 4, #1] Sure, then tomorrow they pass a law that all clinics have to hand over their records."

   B. Certainty they did not have disease

      "[Gp. 2, #5] It's pointless. I've been with the same man and no one else for 10 years. He sure doesn't have the disease. None of us do IV drugs."

<table>
<tr><th colspan="3">I. Reasons given why participants would not get AIDS testing</th></tr>
<tr><th colspan="3">A. Fear of consequences</th></tr>
<tr>
<td>1. Fear of learning they were HIV positive

“[Gp. 7, #2] If I do have it, I don’t want to know.”

“[Gp. 5, #8] If I have it and don’t know it, I live on like a normal person. Why would I want to know I have it? I think denial can be a good thing.”</td>
<td>2. Fear of being seen getting test

“[Gp. 3, #9] Someone sees me in that waiting room and I’m outed.”</td>
<td>3. Fear of records being made public

“[Gp. 1, #10] I don’t believe it when they say it is anonymous. Don’t tell me the testing people don’t say something when they see you later on the streets.”</td>
</tr>
<tr><th colspan="3">B. Certainty they did not have disease</th></tr>
<tr><td colspan="3">“[Gp. 2, #5] It’s pointless. I’ve been with the same man and no one else for 10 years. He sure doesn’t have the disease. None of us do IV drugs.”

“[Gp. 3, #3] I would know it if I was HIV positive. I would be having sores or something.”</td></tr>
</table>

The implementation team should be aware that taxonomies are never exact. Bernard (2006, p. 539) advises researchers not to expect categories to be perfectly “clean.” There is much overlap and indeterminacy in categories.

Matrices are another good way of displaying data. Let us say that in the same study the implementation team collected data on demographic and other personal traits of the focus group participants, and coded these in the texts. These variables could then be compared by participants’ stances on the testing debate, through a matrix. See a hypothetical example below.

Population trait	Participants who said they would not get testing	Participants who said they would get testing
Age	Younger (usually under 35)	Older (usually 35 or over)
Social class	Lower (most under middle class)	Higher (most middle class and above)
Sexual orientation	NA (no substantial differences)	NA (no substantial differences)
Gender	Male (usually)	Female (usually)
Neighborhood	Valley Hill, Orangeville	Rigley Homes (mostly)

The above example is very useful in presenting an easy-to-follow snapshot of the findings. However, the comparisons tend to stereotype the populations. Because this hypothetical information is collected using qualitative strategies only, the population sample that participated in the focus groups would have been purposeful rather than random (as in most quantitative research). Thus findings should not be presented quantitatively, such as in percentages or numbers of participants in each of the cells. The most appropriate way to describe the frequencies in qualitative data would be to avoid use of numbers, but to describe the numbers in words (e.g., “usually,” “over half,” “very few,” “rarely,” etc.).

We will return to this matrix in our discussion of data analysis in mixed methods. A wealth of alternative forms of data displays (such as flow charts and various useful diagrams) can be found in Miles and Huberman (1994).

***Interpretation of findings*.** The final step in any type of data analysis is interpreting the data. This can be done in a variety of ways. Once the implementation team has completed its coding and displays, the team should return to answer the specific questions the research was designed to answer.

If the team implemented the study to identify specific needs, a summary should be made on what these needs are. If the REA was designed to evaluate some intervention, an interpretation should include strengths and weaknesses of the intervention, as well as lessons learned from the intervention process. If the assessment was implemented to effect change in some way, the interpretation should return to this issue and show how the findings suggest need for reform or change in policy. The interpretation can also point out additional questions the findings raise, or show how the results support (or diverge from) past studies on similar topics.

Please now turn to the appendix. At the beginning of the appendix the team will see a transcript of a hypothetical focus group. Read this and then page ahead to the practice exercise for this section where the implementation team will be asked to use topical codes to analyze these data.

The next section will discuss a slightly more complex way of analyzing data.

Using codes as measurement devices

This second level of qualitative analysis described below would be appropriate for times when the implementation team has a fairly weak quantitative section and wants to measure the intensity of some variable (such as intensity of specific attitudes, beliefs, opinions, feelings). Recall that a definition of a variable is simply that—it varies.

Bernard (1995) argues that there are several forms of codes—one that includes information such as the research site and participant, another is the just-discussed topical code, and another is the code that measures. These codes measure the intensity of the topical code. Bernard says this is a “true code” (pp. 193-194).

***Preparing data for analysis*.** The preparation of the data for the codes that measure intensity is the same as it is for simple topical coding.

***Performing the analysis: defining categories*.** Most steps in this level of analysis are the same as they were for the topical coding, except that the implementation team wishes to measure the strength or intensity of the topics that are coded. Recall in the hypothetical study just cited, some focus group participants gave reasons for not wanting to be tested for AIDS. We can know how many times the reasons were cited from simple counting. But particularly in social research settings such as focus groups or interactions documented during observation, occurrences can make the researcher believe that one expressed attitude or stance is much stronger than others, even though it may not explicitly be stated more often. One person may express an opinion and the others in the group applaud the opinion by a clapping of hands or sighs or grunts. For example, participant #9 of group #3 expressed concerns about being “outed” by being seen at an AIDS testing site. His comment may then be met with high fives (which he returns) around the table. The implementation team may then seek a way of defining codes on intensity levels in the codebook, and defining the contexts quite specifically.

A much simpler way to measure intensity is through the words the participants actually state. For example, recall the hypothetical study on STD prevention/treatment needs outlined in the previous section. The example that was provided showed ways to code services relating to viral and bacterial infections. The example is repeated below.

"I believe we need to look at what is really being done in terms of outreach	BAC--syph
to people with syphilis. We had an outreach specialist for that at the clinic years	
ago. But today the emphasis in this area has been strongly on HIV, and people have	VIR-hiv
lost track of the point that these other infections are still out there. It is truly	
my opinion that some of our public health specialists need some serious re-training on	BAC-syph
syphilis and, to some degree, on gonorrhea. I hear all the misinformation. But I don't	BAC-gon
see it happening."	

The person in the above mock interview (whom we assume would be a service provider) is discussing his/her knowledge and opinions about the needs for STD services. The interviewee seems to suggest that the needs for HIV services are being met, but the needs for syphilis and gonorrhea prevention and treatment are not being met. A way of measuring the intensity of expressed needs is to watch for certain words in the texts that relate to intensity, such as "some," "none," "a lot," "very," "strongly," "serious," and on and on. A scale can then be developed as part of the code. In the above example, a subcode could be added to each of the codes designating need, such as BAC-syph/nd (or the code could be rewritten beginning with "need"). Then the implementation team could establish the rules for measuring intensity. The team might specify a list of adverbs and adjectives that will connote "no need" and give this a value of 0, specify a list of adverbs and adjectives that connote "low need" with a value of 1, and so on for "medium need" (value of 2) and "high need" (value of 3).

The coded text might then look like the example that follows. Note that the subcode of “need” receives no value unless some works describing intensity are present.

“I believe we need to look at what is really being done in terms of outreach	BAC—syph/nd
to people with syphilis. We had an outreach specialist for that at the clinic years	
ago. But today the emphasis in this area has been strongly on HIV, and people have	VIR-hiv/nd
lost track of the point that these other infections are still out there. It is truly	
my opinion that some of our public health specialists need some serious re-training	BAC-syph/nd3
on syphilis and, to some degree, on gonorrhea. I hear all the misinformation. But I	BAC-gon/nd2
don’t see it happening.”	

Refining the analysis: Display/presentation of data. Once the coding process is complete, the data can be displayed or presented in any of the formats mentioned under the section on simple topical codes. For example, findings from the above hypothetical study could be presented in a variety of ways. The implementation team might wish to compare the interviewees’ assessment of needs by the kinds of service organizations they represent. A matrix might be useful here (see the following example).

Interviewees' assessment of level of need for treatment/prevention services for some STDs				
Type of service provider of interviewee expressing assessment	**HIV/AIDS Needs**	**Syphilis Needs**	**Gonorrhea Needs**	**Chlamydia Needs**
	Assessment			
Healthcare institution	Medium	High	High	High
STD social service agency	Medium	High	High	Medium
Correctional institution	High	High	High	None
Home health services	High	Low	None	High

Here again, the implementation team should avoid presenting these findings in quantitative terms (such as use of percentages and means) unless the study was done with a random sample.

Interpretation of findings. Interpretation would involve all the issues discussed under this heading in the section on simple topical codes.

***Please now turn to the appendix*.** At the beginning of the appendix you will see a transcript of a hypothetical focus group. Read this and then page ahead to the practice exercise for this section where the implementation team will be asked to use a combination of codes to analyze these data.

The next section will discuss a considerably more complex way of analyzing data.

Analysis of emergent research sequence

Early in this chapter we issued a word of caution about categorizing data—that it ultimately involved subjectivity on the researcher's part and may not actually reflect the way the target community categorizes their experiences. In some cases, as we mentioned, classification schemes are understood nearly universally. In other cases the researchers need to try very hard to understand the categories of meanings that those studied apply to their experiences. There is no way that this can be done with perfect precision. Even in the most tightly knit cultural groups, complex individual differences abound. Furthermore, no researcher can maintain absolute objectivity. Every research strategy outlined in this curriculum, as well as all forms of research strategies known worldwide, involve some level of subjectivity in planning and implementation. However, there are always ways to reduce the subjectivity, get the target community more involved in research decisions, and thus reduce the risk of researcher bias emerging in the results. James Spradley (1980) developed a sequence for conducting ethnographic research that reduces these risks to some extent. In what we will term his "emergent research sequence," Spradley provides complex but specific rules on how to develop taxonomies and identify themes or patterns. The steps in the sequence take a 195-page book to describe, but we will only deal with some of his high points of data analysis here.

The data analysis techniques derived from Spradley's model would be valuable if the implementation team is relying on qualitative research as the predominant methodology. The techniques would be particularly valuable when all or some part of the study is focusing on previously unexplored topics or where researchers really need to understand the meanings that the target community attribute to phenomena.

Preparing data for analysis. The implementation team should transcribe all audio- and/or videotaped data when collected and (where possible) type hand-written notes from interviews, focus groups, and observation. This follows the steps outlined under this heading for simple topical coding.

However, Spradley argues that the categorization of data should be concurrent with data collection and not begin at a later stage. He maintains that there are certain relationships that are nearly universally understood from which taxonomies could be formed, and some or all of these relationships should be explored during the earliest stages of data collection.

The implementation team was introduced to these relationships in the chapter on "Observation" in the section, Data Collection, Qualitative Strategies (see below). The hypothetical example had as the focus of the study a particular organization that recently experienced an unexplained outbreak of a certain STD among its employees. During observation, the researchers would probably have organized observation guides to explore one or more of these relationships. One suggested was the "location-for-action" relationship where observers were considering the possibility that some sexual contact was occurring during working hours, and would collect data on rooms where intimate relations were possible.

Spradley's "semantic relationships" are repeated below (1980, p. 93).

1. Strict inclusion	X is a kind of Y
2. Spatial	X is a place in Y X is a part of Y
3. Cause-effect	X is a result of Y
4. Rationale	X is a reason for doing Y
5. Location-for-action	X is a place for doing Y
6. Function	X is used for Y
7. Means-end	X is a way to do Y
8. Sequence	X is a step (stage) in Y
9. Attribution	X is an attribution (characteristic) of Y

***Performing the analysis: Defining categories*.** Spradley does not use the term "code" in his work. But the relationships he is describing are roughly equivalent to what we have described as the coding process, where "Y" in the semantic relationships might be a topical code followed by the label of the relationships (e.g., "function," "spatial") and "x" might be the subcode[s]. The difference is that here the researcher is following a very specific process for categorizing the data—one Spradley suggests embraces relatively universal relationships. The researcher is not creating his/her own categories.

Let us return to the hypothetical study. Because the researchers (and their initial informants) in our "study" really have no idea why this outbreak has taken place within the organization, researchers might explore a number of possibilities simultaneously. In addition to observation, they might send out a confidential questionnaire asking open-ended questions. Two questions might be the following.

A. Have you ever had sexual contact with anyone in this organization?
   1. Yes
   2. No [YOU ARE NOW FINISHED WITH THIS QUESTIONNNAIRE. PLEASE PUT THE QUESTIONNAIRE IN THE PRE-SEALED ENVELOPE AND DROP IT IN THE SLOT IN THE SECURED SURVEY BOX IN THE UTILITY ROOM]

B. Please describe the circumstances under which you have had sexual contact with someone in this organization (e.g., when, where, why, relationship to person, STD prevention strategies used). If you have had sexual contact with more than one person, please describe each situation separately.

When the researchers are categorizing (coding) the data from the questionnaires, a number of semantic relationships can be explored. Some might be "x is a place for having sex with people in the organization," "x is a way to have sex with people in the organization," or "x is a step in having sex with people in the organization (which may or may not include STD preventive strategies)." Another relationship that delves deeper into meanings is "x is a reason for having sex with people in the organization." Imagine that the list of possibilities for the latter relationship looks like the following:

A. Reasons for having sex with people that work in this organization.

1. Looks (the sexual partner is attractive)
2. Status (the sexual partner has high status)
3. Hot pursuit (the sexual partner pursued the respondent tenaciously)
4. Convenient place (the workplace/organization was a convenient place to have sex)
5. General desire (the respondent is just plain sexually attracted to others)
6. Workplace tolerance (people at work are open to homosexual/other alternative relationships)
7. Workplace conformity ("everyone else" is sexually involved at work)
8. Coercion (sexual partner coerced through threat of physical force or workplace sanctions)
9. Easy to conceal (workplace affairs can be concealed from outside partners because they take place during work, etc.)

Now of course the researchers would be developing the above list themselves and some subjectivity can come into play while generating the subcategories. But then something interesting happens. The researchers are then asked to search for "dimensions of contrast" within the list (which is actually a taxonomy), and to seek the help of the target community in doing this[3] . In our hypothetical study, the researchers would request the assistance of a sample of the organization's employees to search for these contrasts. One way that Spradley recommends doing this is by writing the subcategories (in this case the "reasons") on cards and asking oneself and members of the target community to go through each card and attempt to determine how the current card is different from the last. Ultimately piles of similar "reasons" are formed from the process. Like most forms of categorization, the piles will never be "clean." Overlap will occur. The researchers then keep a list of the contrasting terms that have been named, making note of the contrast terms that come up most consistently.

Let us say in our hypothetical study that the consistent dimensions of contrast for the "reasons to have sex with people in the organization" are the following:

1. Sexual attraction
   a. Reason requires sexual attraction
   b. Reason does not require sexual attraction
   c. Reason apparently requires sexual attraction, but also requires something else
2. Workforce environment
   a. Reason requires something in the workforce environment
   b. Reason does not require something in the workforce environment
3. Pressure of sexual partner
   a. Reason requires strong pressure by sexual partner
   b. Reason does not require strong pressure by sexual partner

Let us also say that some of these dimensions of contrast came up in the majority of other taxonomies of semantic relationships. For example, "workforce environment" might come up as a dimension of contrast while coding the data for the relationships "x is a place for having sex with people in the organization" or "x is a way to have sex with people in the organization." If this is the case, then a clear pattern or theme has emerged.

Please read the following example from the research of Jill Florence Lackey to see how important it is to involve the target community in qualitative research analysis. Following this is a section that suggests strategies for tying patterns or themes back to the research question.

I was studying “the homeless” when the term first began emerging in the media. At the time I began my studies of this population, very little research had been done on the topic. It was pretty much unexplored territory. I was conducting interviews and participant observation at various soup kitchens in an urban center and was attempting to build a taxonomy on types of “homeless,” from the insiders’ perspective. During interviews I noted that some of the long-time homeless also used the term “street people” for those that frequented the soup kitchen. Thus, as I began coding my data I did not distinguish between the term “homeless” and “street people.” Ultimately I developed a list of twenty subcategories of “homeless/street people.” I wrote each type on cards and then took the cards to a sample of nine key informants (long-time homeless) and began to ask them to sort the cards into piles. The first three informants seemed confused by my directions and refused to continue the exercise. When I got to the fourth the dialogue went something like the following.

“I have here a list of twenty kinds of homeless (or street people) that you and others at the meal site described during the interviews. Each kind is written on one of the cards. I want you to look at each card and try to figure out what makes this kind of homeless or street person different than the kind in the next card. You may sort them into piles when you see similarities.”

The man began to look through the cards. “They ain’t the same, you know.”
“What isn’t?”
“Homeless and street people. They ain’t the same.”

The gentleman patiently explained to me that street people weren’t necessarily homeless, nor was every homeless person necessarily known by the soup kitchen crowd as a street person. “Some of the homeless you don’t even see,” he curiously remarked.

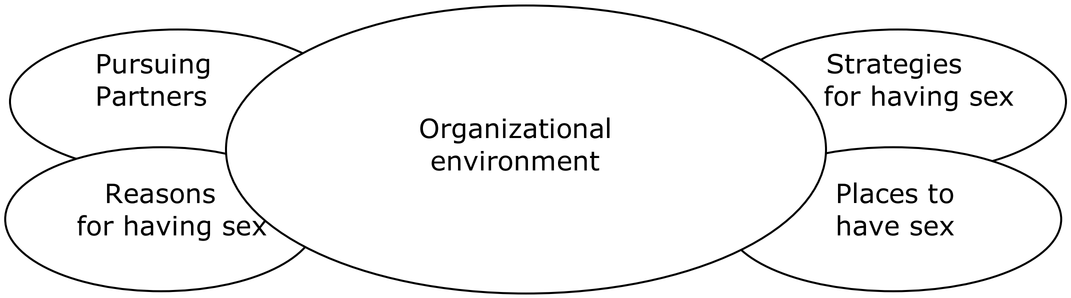
I went back to the drawing board and developed a new coding system, coding separately for the terms “homeless” and “street people.” After going through the entire painstaking process again, I and the soup kitchen informants eventually identified an overarching theme for “street people.” I labeled it “public living” (a theme that clearly did not emerge for the “homeless”). My informants concurred and I then understood what the gentleman meant.

The implementation team should note that critics of the type of analysis just presented argue that it carries the assumption that people think in more or less binary terms—and there is really no way that one can know this.

***Refining the analysis: Display/presentation of data*.** Using Spradley's strategies, once patterns have been identified through analysis, data can be presented in any of the ways that have been summarized in this chapter. Because this approach to data analysis is designed specifically for ethnographies, findings tend to be holistic in nature and the implementation team would have to present findings in ways that related back to the research question. In the case of the organization trying to understand why they had the outbreak of an STD, patterns might emerge on any topic (e.g., departmental differences, power relations, pay scales). These other topics would be most interesting to explore, but because of the time concern in the REA, researchers should focus the findings on the central question. One way that the implementation team might present the findings on this hypothetical study is the following.

In every aspect of our study of the Ace Agency, findings suggested an organizational culture that accentuated unrestrained sexual contact, made the contact easy to occur, or at minimum did little to hamper it. When we conducted a confidential open-ended survey of employees, over one-quarter reported having sexual contact with at least one person in the organization. When these participants described their sexual encounters, the theme of organizational environment came up more often than any other.

[Here the researchers should provide a series of interview quotes in support of the graphic below. The graphic should summarize some of the major patterns.]



[Here the researchers should present taxonomies, matrices, or other graphic displays (if useful) that summarize major themes.]

A survey question specifically asked respondents to relate any form of STD prevention they had taken during sexual contact with someone from the organization. However, less than half of the 25 respondents who said they had sex with someone from the organization even mentioned the topic, and of these, only two described taking preventive measures.

Interpretation of findings. Interpretation would involve all the issues discussed under this heading in the section, "Simple topical codes". In the case of the above hypothetical study, the implementation team might make recommendations (based on the research) on how to change the organizational culture and environment to reduce the transmission of the STD. However, these recommendations would ideally be negotiated with the target community to avoid researcher bias.

Please now turn to the appendix. At the beginning of the appendix the team will see a transcript of a hypothetical focus group. Read this and then page ahead to the practice exercise for this section where the implementation team will be asked to use some of these strategies to analyze the data.

Getting started

As in data collection, the implementation team can select from three levels of difficulty. However, the level of difficulty should be chosen based on the data analysis strategy that bests suits the team's research design, the purpose of the study, and the data being collected--not on the difficulty factor.

***Lowest level of difficulty*.** Here the implementation team could select the analysis strategies summarized under "simple topical coding." The team might also try and develop at minimum one visual display, such as a taxonomy table or matrix, as this may be important for later presentation of findings. If this option is chosen, begin with the practice exercise in the appendix.

Medium level of difficulty. Here the implementation team could combine topical coding with measuring codes, if this fits the study questions. The team should also try and develop at minimum one visual display, such as a taxonomy table or matrix, as this can be valuable later in presentation of findings. If this option is chosen, begin with the practice exercises in the appendix.

Highest level of difficulty. Here the implementation team could follow the analysis steps in the "emergent research sequence." The team should also try and develop at minimum one visual display, such as a taxonomy table or matrix. If this option best fits the study questions, begin with the practice exercise in the appendix.

[1]Other mainstream specialists in qualitative data analysis (not necessarily ethnographers) also emphasize the inductive approach (e.g., Lincoln & Guba, 1985; Strauss & Corbin, 1998).

[2]Most researchers edit texts slightly for the sake of the audience. The use of brackets is one good strategy.

[3] We do recognize that in studies involving STDs that confidentiality/anonymity issues may emerge that may make it impossible to follow-up with members of the target community after data have been collected.

Section 7, Chapter 2: Qualitative data analysis

7.2.3 Learning activities

Time to review

The implementation team should now ask each other the following questions.

1. What are the basic differences between the three forms of qualitative data analysis presented in this chapter?
2. How are the data prepared for analysis in qualitative research?
3. What are the ways that qualitative data analysis is performed in topical coding?
4. What are the ways analysis is performed when using measurement codes?
5. What are the ways analysis is performed when using the emergent research sequence?
6. What are some of the ways data analysis can be displayed?
7. What is the difference between display or presentation of data and interpreting findings?

Analyzing qualitative data systematically

Once decisions have been made about the way data will be analyzed, the implementation team should respond to the following questions to check for consistency (also see more detailed worksheets in the appendix).

ANALYZING QUALITATIVE DATA SYSTEMATICALLY

1. Will the team have more than one person involved in coding?
2. Did team members who will be involved in coding generally agree on codes for the practice exercise[s]?
3. Do the data analysts have a plan for organizing the codebook and keeping a log of the texts that are coded?
4. Are there key informants in the study that the team can call upon to, at minimum, review any patterns that the data analysts discover?
5. Do the data analysts have a start and an end date for qualitative analysis?

The team is now ready to begin analysis.

Quality control: Checking progress

Once the data analysis is underway, the implementation team should perform quality checks on the analysis at agreed-upon intervals. The researchers can accomplish this by responding to a series of questions. (The more detailed worksheets are printed at the end of this chapter.)

QUALITY CONTROL ASSESSMENT: QUALITATIVE DATA ANALYSIS

1. Did the data analysts reach consensus on the initial codes or semantic relationships?
2. Have the analysts reached consensus on any taxonomies developed?
3. Have the analysts reached consensus on any emerging patterns?
4. Have any key informants corroborated these patterns?
5. Will the analysts have the texts (from interviews, focus groups, life histories, or observation) to illustrate these patterns?
6. Have the data analysts found ways to organize the data into clear but visually interesting forms?

7.2.4 Resources

Chapter references

Beebe, J. (2001). *Rapid assessment process: An introduction*. Walnut Creek, CA: AltaMira.

Bernard, H.R. (2006). *Research methods in anthropology: Qualitative and quantitative approaches* (4th ed.). Lanham, MD: AltaMira.

Bernard, H.R. (1995). *Research methods in anthropology: Qualitative and quantitative approaches* (2nd ed.). Walnut Creek, CA: AltaMira.

Lincoln, Y., & Guba, E. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage.

Creswell, J.W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches* (2nd ed.). Thousand Oaks, CA: Sage.

Miles, M.B., & Huberman, A.M. (1994). *Qualitative data analysis* (2nd ed.). Thousand Oaks, CA: Sage.

Pelto, P.J. & Pelto, G.H. (1987). *Anthropological research: The structure of inquiry* (2nd ed.). Cambridge: London.

Spradley, J.P. (1980). *Participant observation*. New York: Holt, Rinehart, and Winston.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd ed.). Thousand Oaks, CA: Sage.

Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches*. Thousand Oaks, CA: Sage.

Additional resources on qualitative data analysis

Tesch, R. (1990). *Qualitative research: Analysis types and soft-ware tools*. New York: Falmer.

Denzin, N.K., & Lincoln, Y.S. (Eds.). (2005). *The SAGE handbook of qualitative research* (3rd ed.). Thousand Oaks, CA: Sage.

Glaser, B.G. & Strauss, A.L. (1967). *The Discovery of Grounded Theory*. Thousand Oaks, CA: Sage.

Section 7, Chapter 2: Appendix

HYPOTHETICAL FOCUS GROUP

<u>Situation:</u> *A series of focus groups is underway to assess the need for a program that would educate parents on the best ways to talk to their children about STDs. The facilitator has already defined STDs and "sex.." Below is an excerpt from one focus group at a community center. These parents do not know each other, but all 10 are mothers, live within a specific geographical area, and all have children between the ages of 13 and 15.*

<u># 4:</u> "I'm interested in talking to my daughter about sexually transmitted diseases, and I have to some extent, but I really hope this will never be an issue in her life. I have to admit that I feel pretty strongly about abstinence [several others around the room reply "amen"] and I guess for that reason I haven't considered many other options.

<u>#1</u>: I agree with #4 here, but just to be on the safe side, I intend to talk to my daughter about condoms if there ever is a time I think she is about to have sex—hopefully that will be a long time off.

<u>#10</u>: I've already gone through the bit on condoms with my son. When he has sex I don't want anyone coming and suing me because he gave some little girl AIDS or got her pregnant.

<u>#2</u>: Maybe if I didn't have all girls . . .

<u>#9</u>: I tell my son that he should watch out on who his partner is. He should know something about her . . .

<u>#10</u>: Nah, that don't mean anything . . .

<u>#1</u>: She's right—you can't know if someone has a disease or not.

<u>#3</u>: I'd go with advising my son to have his girlfriends tested [negative grumbles can be heard around the room]. Of course, he'd never do it.

<u>#7</u>: None of them would do it.

<u>#5</u>: They'd both have to do it. I think it's a good idea, but it's too complicated. When do you go in for testing? Every time my son takes out a new girlfriend? Maybe the girlfriend is messing around. It's too complicated.

#6: This is way too complicated. See where it's leading? Abstinence is the only answer. If I started talking to my daughters about preventing diseases, they'd start to think that I was okay with them having sex before marriage and maybe they'd start to experiment. Abstinence is really the only answer.

#7: I have a daughter and a son. I wish they would both hold off, but if I started talking to my kids about abstinence (and my son has probably already done his share of fooling around), they might look at me like I'm nuts. I tell them about everything that can happen to them if they catch a disease—how they can die from AIDS, end up sterile or insane from some of the others, get sores all over the place, how you can't get rid of some of the diseases. I tell them that hardly anything works to prevent these things once they start having sex.

#10: You can tell them all that and I do, but you've gotta tell them about condoms too. Condoms will protect them from most of the diseases. You really need to do that.

FACILITATOR: #8, you haven't said anything. What is your strategy?

#8: I've just been listening to all of this. I'm not sure what the answer is. I heard you can get some STDs from kissing. I think if you are going to be really sure of everything, you've got to avoid contact with all body fluids. I just don't know.

#7: How do you do that?

#8: I tell my daughter that if she ever starts feeling hot and heavy before she's legally married, she should fool around with the clothes on, maybe do a little rubbing—you get what I mean? And don't do any really hot necking.

#9: I told my son to do that too. I doubt that he does though.

#2: I never thought of that. I guess it's kind of a way for girls to remain virgins also—well until they're married. It might be worth a try.

#5: So you're going to teach your daughters just how to rub a guy so he has an orgasm with his clothes on, and vice versa. And the guy is just going to go along with this?

#8: Well, I'm assuming that they [her girls] haven't tried it out yet. It probably isn't the best idea.

#7: It's weird.

#10: It'll never work.

#6: You just gotta go with no sexual contact at all. Abstinence is the only answer.

#5: I'd never convince my son of that. I guess condoms is the best of all, but it's not perfect.

#10: Abstinence is never going to be the thing in the world we live in, so condoms is the best way. I'll bet all your kids have already had sex and you don't even know it. Your daughters aren't what you think they are. At least I know my son will tell me when he has sex.

#2: My daughters would too, if they ever had sex.

#7: "If" is the word. I mean we've got to tell them something. My "scare tactics" have worked. My son and daughter talk about that all the time. I see them watching TV and they say, "Look at that guy sticking his tongue in her mouth. She could have herpes." Or I'll hear them say something when they have this hot sex scene, like, "How do they know the other one doesn't have AIDS?" It works.

#4: I guess I could try it, but there's something that seems somehow wrong with "scare tactics."

[THE DISCUSSION MOVES TO ANOTHER TOPIC]

WORKSHEET CHAPTER 2A: PRACTICE EXERCISE WITH SIMPLE TOPICAL CODING

At least two members of the implementation team should be involved in this exercise. The team should not expect to have "absolutely correct" answers while doing this exercise. There are a number of ways that the text can be coded and analyzed. The trick is to look for agreement among coders.

1. Working independently, each member of the team should read the transcript of the hypothetical focus group through once. Each person should look for one or more "umbrella" codes. When team members have completed these, they should identify subcodes.   Label the codes and mark these down in a mini codebook, defining the codes and subcodes as well as possible. Realize that no categories are ever perfectly "clean." Team members will always have to deal with some overlap and ambiguity. And independent coders will rarely come up with the same labels and definitions.
2. Still working independently, team members should mark the codes and any other remarks they may wish to add in the right margins of the mock transcript. Team members should also keep a log (of individual choosing) of the text excerpts they have coded.
3. The team members should develop at least one taxonomy from the categories.

4. The independent coders should now come together and compare their taxonomies. If the codes selected and the initial taxonomies appear similar, the team should reach a compromise on the final forms. If they appear very different, the team needs to discuss the differences and try the exercise again.
5. The team should attempt to develop a matrix from any of the data that have been analyzed.

See some possibilities for analyzing this text with topical coding at the end of this section. But the team should not look at these examples before completing their own exercise. Recall, however, that there are no "perfectly correct" answers.

WORKSHEET CHAPTER 2B: PRACTICE EXERCISE USING CODES AS MEASUREMENT DEVICES

At least two members of the implementation team should be involved in this exercise. The team should not expect to have "absolutely correct" answers while doing this exercise. The trick is to look for agreement among coders.

1. If the team has not yet completed the exercise on topical coding, at least two team members should do this now.
2. Working independently, the team members should look over the codes and taxonomies and decide which categories should and could be measured for intensity of responses. Select at least one taxonomy to measure intensity of subcategories.
3. Select values for the levels of intensity (such as 1-3 or 1-5). Team members should specify the textual cues/indicators that will define the level of intensity (such as specific adverbs and adjectives, reservations about a particular response, reassertion of a point, retracting a point, etc.) Record this information in the codebook.
4. Mark the codes with the measurements in the right margin of the text. In this case the "unit of analysis" will be individual focus group participants, which means that the intensity measure for each code should be applied once to each participant in the focus group. [Hint: it is easiest to go through all the comments of one participant first, then decide on one value for each code defined for that participant.)

5. The independent coders should now come together and compare their results. Did the team members generally agree on the kinds of codes that could be measured for intensity? Were the definitions for the levels of intensity similar? If the results were similar, the team should compromise on the final forms. If they appear very different, the team should discuss the differences and try the exercise again. Recall that when conducting the actual study, the initial coding the team decides on will come from a sample of the data and that coding will then be applied to all the data (undoubtedly with some modifications in process).

See some possibilities for analyzing this text using measurement codes at the end of this section. But the team should not look at these examples before completing their own exercise. Recall, however, that there are no "perfectly correct" answers.

WORKSHEET CHAPTER 2C: PRACTICE EXERCISE WITH THE EMERGENT RESEARCH SEQUENCE

At least two members of the implementation team should be involved in this exercise. Again, the team should not expect to have "absolutely correct" answers while doing this exercise. The trick is to look for agreement among coders.

1. Working independently, each member of the team should read the transcript of the hypothetical focus group through once. Each person should list as many "semantic relationships" as possible that could be explored in the text. The team members should then come together and see if they agree on the semantic relationships they could explore, should they choose to develop the lists.
2. Working independently again, team members should form a list from the semantic relationship of "means-end" and complete a taxonomy of "x is a way mothers could advise their adolescent children on preventing STDs" (ignore gender of children for this exercise). This list should look very much like (or exactly like) the taxonomy that was developed in the previous exercises. Each team member should then list the variables on cards and role-play that each is a key informant for the study (i.e., a focus group participant). Each team member should then form piles of similar variables and identify at least one dimension of contrast among the variables in the taxonomy, "ways mothers could advise their adolescent children on preventing STDs."
3. The team members should come together and see if they agree generally on the way they have formed their piles and on at least one dimension of contrast. Where modest differences occur they should reach a compromise. If the differences appear very different, the team should discuss the differences and try the exercise again

See some possibilities for analyzing this text using this model at the end of this section. But the team should not look at these examples before completing their own exercise. Recall, however, that there are no "perfectly correct" answers.

POSSIBLE WAYS OF ANALYZING THE HYPOTHETICAL FOCUS GROUP USING SIMPLE TOPICAL CODING

A. Ways mothers could advise adolescent children on preventing STDs
   1. Maintaining abstinence
   2. Using condoms
   3. Knowing something about the sexual partner[s]
   4. Telling children the negative consequences of STDs
   5. Avoiding contact with bodily fluids
   6. Having sexual partners tested for STDs

B. Ways mothers of adolescent daughters said they advised children on preventing STDs
   1. Maintaining abstinence
   2. Telling daughters the negative consequences of STDs
   3. Avoiding contact with bodily fluids

C. Ways mothers of adolescent boys said they advised children on prevention of STDs
   1. Using condoms
   2. Telling sons the negative consequences of STDs
   3. Knowing something about the partner[s]
   4. Having partners tested for STDs

D. Assumptions mothers of adolescent boys had of sexual activity of children
   1. Son has already had sex
   2. Son will probably have sex before marriage

E. Assumptions mothers of adolescent girls had of sexual activity of children
    1. Daughter has not already had sex
    2. Daughter will probably not have sex before marriage

(If the implementation team was at minimum able to identify most of the ways the parents said they advised their children on STD prevention, and some of the differences being expressed by parents of boys and girls, the team did well.)

Ordinarily a matrix would not be created on findings from one focus group, but if one imagines that these findings are based on multiple focus groups, the matrix might look like the following.

Mothers' STD Topics	Male children	Female children
Assumptions mothers made on sexual activity of adolescent children	o Son has already had sex o Son will probably have sex before marriage	o Daughter has not already had sex o Daughter will probably not have sex before marriage
Advice mothers reported giving adolescent children on preventing STDs	o Using condoms o Telling sons the negative consequences of STDs o Knowing something about the partner[s] o Having sexual partners tested for STDs	o Maintaining abstinence o Telling daughters the negative consequences of STDs o Avoiding contact with bodily fluids

POSSIBLE WAYS OF ANALYZING THE HYPOTHETICAL FOCUS GROUPS USING CODING AS A MEASUREMENT DEVICE

Intensity of responses could best be measured in the following taxonomy.

A. Ways parents could advise their adolescent children on preventing STDs
1. Maintaining abstinence
2. Using condoms
3. Knowing something about the sexual partner[s]
4. Telling children the negative consequences of STDs
5. Avoiding contact with bodily fluids
6. Having sexual partners tested for STDs

The above could also include subcodes for parents of girls and parents of boys.

- An example of a label for a measurement code could be
  ADV-abst1 (for advice-abstinence/intensity level of 1)

- Another example could be
  ADV-pg/abst1 (for advice-parent of girl[s]/abstinence/intensity level of 1)

One way the team could have measured the level of intensity appears below.

Levels of intensity					
Level	Strongly against (advice)	Against (advice)	Neutral	For (advice)	Strongly for (advice)
Value of code	1	2	3	4	5
Indicators of level	OTerms describing strength of opinion, such as *never, absolutely, really, very, no answer, worst, etc.* O Participant restates opinion later (without reservations) O Participant gives multiple reasons for opinion, or examples where advice would not work O Participant ridicules idea	O Participant suggests is *against* advice, but texts lacks terms describing strength of opinion O Participant suggests is *against* advice, but with *some* reservations	O Participant states neutrality O Participant sees arguments both ways O Participants states is *for* or *against* advice, but states *strong* or *multiple* reservations O Participant later retracts opinion (but does not specifically change opinion)	O Participant suggests is *for* advice, but text lacks terms describing strength of opinion O participant suggests is *for* advice, but with *some* reservations	O Terms describing strength of opinion, such as *always*, *absolutely*, *really*, *very*, *no answer*, *only answer*, *best*, *etc.* O Participant restates opinion later (without reservations) O Participant gives multiple reasons for opinion, or examples where advice works

Once the analysis is complete, one would have little to report back on from one focus group, but if the other focus groups followed the patterns of this hypothetical focus group, then a description of results might look like the following.

In general, mothers who said they preferred "abstinence" as a way to advise their children on STD prevention expressed very strong opinions on the topic. See focus group excerpts below.

[Researchers would then add a series of quotes]

On the other hand, mothers appeared polarized by gender of their children when discussing abstinence. We found most of the mothers expressing strong opinions on abstinence were mothers of girls, while mothers of boys were often strongly opposed to the strategy. See examples that follow.

[Researchers would then add a series of quotes]

POSSIBLE WAYS OF ANALYZING THE HYPOTHETICAL FOCUS GROUPS USING THE EMERGENT RESEARCH SEQUENCE

Possible semantic relationships that could be explored in the focus group text are:

Rationale: X is a reason for talking to adolescent about sexual risk taking

Rationale: X is a reason for not talking to adolescent about sexual risk taking

Cause/effect: X is a result of talking to adolescent about sexual risk taking

Means-end: X is a way for adolescents to remain virgins

Means-end: X is a way for mothers to advise their adolescents on preventing STDs

The relationships could also mention gender of child being discussed.

The possible taxonomy for the latter relationship follows:

A. Ways mothers could advise their adolescent children on preventing STDs

1. Maintaining abstinence
2. Using condoms
3. Knowing something about the sexual partner[s]
4. Telling children the negative consequences of STDs
5. Avoiding contact with bodily fluids
6. Having sexual partners tested for STDs

Dimensions of contrast

One possible dimension of contrast among variables in the above taxonomy comes under the heading of "sexual activity." See below.

Sexual Activity			
	Method implies that adolescent is having direct sexual activity	Method implies that adolescent not having direct sexual activity	N/A: Method does not make implications either way
Ways mothers could advise their adolescents on preventing STDs	o Using condoms o Knowing something about the sexual partner o Having sex partners tested for STDs	o Maintaining abstinence o Avoiding contact with bodily fluids	o Telling children the negative consequences of STDs

Does the question about whether the prevention methods imply direct sexual activity get to the deeper "meaning" of the focus group debates?

WORKSHEET CHAPTER 2D: ANALYZING QUALITATIVE DATA SYSTEMATICALLY

1. Will the team have more than one person involved in coding?

Yes _______ No _______

It is very important to have more than one person involved in coding. If no one else from the team is available, try and involve someone from the target community or the collaborating stakeholders. Of course they will need to use the curriculum as well.

---

1. Did team members who will be involved in coding generally agree on codes for the practice exercise[s]?

Yes _______ No _______

If no, try the exercise[s] again.

---

2. Do the data analysts have a plan for organizing the codebook and keeping a log of the texts that are coded?

Yes _______ No _______

If no, now is the time to develop this plan.

3. Are there key informants in the study that the team can call upon to, at minimum, review any patterns that the data analysts discover.

Yes _______ No _______

This is strongly recommended, but it is not always possible, particularly if the team cannot keep names of informants, even in coded forms.

4. Do the data analysts have a start and an end date for qualitative analysis?

Yes _______ No _______

If no, now is the time to set the dates.

Start date ___________________ Stop date _______________________

WORKSHEET CHAPTER 2E: QUALITY CONTROL ASSESSMENT: QUALITATIVE DATA ANALYSIS

1. Did the data analysts reach consensus on the initial codes or semantic relationships?

Yes ______ No ______

If yes, indicate below what these initial codes or semantic relationships are (list only the relationship title for the semantic relationships--e.g., function, means-end).

If no, indicate below what the team does agree on (if anything) and set a future time to revisit this issue.

2. Have the analysts reached consensus on any taxonomies that have been developed?

Yes ______ No ______

If yes, list below the titles of the taxonomies.

If no, indicate below what the team does agree on (if anything) and set a future time to revisit this issue.

3. Have the analysts reached consensus on any emerging patterns?

Yes ______ No ______

If yes, briefly describe the patterns below.

If no, indicate below what the team does agree on (if anything) and set a future time to revisit this issue.

4. Have any key informants corroborated these patterns?

Yes ______ No ______

This is strongly recommended, but it is not always possible, particularly if the team cannot keep names of informants, even in coded forms.

5. Will the analysts have the texts (from interviews, focus groups, life histories, or observation) to illustrate the patterns?

Yes ______ No ______

This is a very important aspect of qualitative research. The team needs to do this.

6. Have the data analysts found ways to organize the data into clear but visually interesting forms?

Yes ______ No ______

If yes, briefly describe the forms below (e.g., charts, matrices).

If no, set a future time to revisit this issue.

Section 7, Chapter 3: Quantitative data analysis

7.3.1 Intended learning outcomes

The intended learning outcomes of this chapter on quantitative data analysis follow.

At the end of this chapter the implementation team will be able to:

1. Understand basic procedures of quantitative data analysis common in rapid assessments;
2. Know when to use a particular data analytic technique; and
3. Be able to provide straightforward interpretation of data analytic results.

Section 7, Chapter 3: Quantitative data analysis

7.3.2 Introduction

The topic of quantitative data analysis can be highly complex. However, this curriculum is designed to introduce the implementation team to a few strategies that can be implemented without use of statistical software programs or courses in statistics. For more complex procedures the implementation team might consider consulting with someone with professional background in this area, or the team can check out the addendum at the end of this section for more sophisticated forms of quantitative data analysis.

Sahai and Khurshid (2002) define data analysis as the process of reducing accumulated data to a manageable size, developing summaries, looking for patterns, and performing statistical analysis (pg. 74). Quantitative data – often called numerical data – are data obtained by using numerical measurement. Numerical data involve either continuous measurements or counts (pg. 181).

The main aim of quantitative data analysis is to summarize numerical information so that relevant information emerges. The summaries of quantitative data can take the form of simple percentages and averages or complex statistics involving relationships between variables. This chapter will focus on several basic analysis procedures – from low level of difficulty (calculating a percentage of "yes" responses to a single survey question, calculating an average from a count variable, and comparing averages between two groups) to medium level of difficulty (constructing a crosstabulation of two nominal variables) to high level of difficulty (performing a statistical test to see if one group has responded "yes" to a question significantly more than another group). Although we state three levels of difficulty, the highest level difficulty example discussed in this chapter is still presented at a level for beginning data analysts. More complex data analysis techniques exist, of course, but what is covered here should be sufficient for most people using this curriculum. Implementation team members wishing more information on quantitative data analysis are referred to the table of statistical analysis software and other resources located in the appendix.

Terms. Quantitative data analysis involves use of some terms.

- A "statistic" is a numerical value used as a summary measure for a sample. Some examples are the sample mean (also called the arithmetic average) and the sample standard deviation.
- A "variable" was defined in the previous chapter as "any of the elements, concepts, or categories in theoretical propositions that are thought to vary in degree or in kind" (Pelto & Pelto, 1987, p. 142). Or it can be defined as any quantity that varies--an aspect or characteristic of a person, object, or situation that can assume different values.
- A "continuous variable" is a numerical variable in which the values can be any number along a continuum. Height is an example. Although a person's height is typically recorded to the nearest inch or half-inch, theoretically, height could be measured to the nearest tenth of an inch or even more precisely.
- A "count variable" is a numerical variable in which measurements only take on whole number values. An example is the number of teenagers living in a household. Outcomes are likely to be 0, 1, 2, etc., but values between these whole numbers are not possible.
- A "nominal variable" is a type of categorical variable in which the categories have meaning by name only. It can be a dichotomy (either "yes" or "no") or a set of categories. For example, the variable political affiliation may be categorized by the names Republican, Democrat, and Independent, but these categories have no numerical meaning. We can, however, calculate the percentage of participants in each category, and that will provide us with useful quantitative information like which political affiliation is most common in a sample and how much more common is it compared to other categories.
- An "ordinal variable" is a type of categorical variable in which the categories of the variable increase in magnitude as one reads across the categories (i.e., they are ordered). An example of an ordinal variable commonly used in survey work is a "Likert scale". Respondents specify their level of agreement to a statement. An example of a Likert scale is shown below:

Sex education should be part of every high school's curriculum. [CIRCLE ONE]:

1	2	3	4	5
Strongly disagree	Disagree	Neutral	Agree	Strongly agree

The differences between the categories of a Likert scale (or any ordinal variable) may not be of the same magnitude. Generally, the researcher assumes these differences are the same, though, and it is considered acceptable to analyze a Likert scale as if it were a regular continuous interval variable.

Advantages and limitations of quantitative data analysis

***Advantages*.** The results of numerical analysis are considered unambiguous in the sense that communicating the results in the form of numbers is accepted and understood across disciplines and cultures, whereas qualitative analysis requires categorization of observations often in a more subjective manner.

It is also much easier to concisely summarize quantitative information than qualitative data. Quantitative data provide good estimates of a sample's attitudes or behaviors (albeit, not cultural meanings). Quantitative data are best used when concise description of populations as samples are desirable, especially with large sample sizes that make it difficult to describe qualitatively.

***Limitations*.** Many numerical quantities are difficult to compute – often requiring the use of a statistical software program or at least a hand-calculator. With large volumes of information, use of computer databases and statistical packages are essential. Skill at entering data into the computer is also necessary when working with large sets of data.

Because of the abundance of numerical statistics, it is easy for a researcher to intentionally or unintentionally communicate results in a misleading way. Even when presenting a graph – the way the graph is displayed can vary from one analyst to the next, given the same information.

In addition, quantitative data tend to reflect only predetermined variables and response options. Thus they can confirm but not generate hypotheses or hunches. Qualitative approaches are more open to generating new hypotheses.

A word about computer programs

Although the data analysis problems presented in this chapter can be calculated by hand (mainly with a basic hand-calculator), we realize there is likely to be an occasion when the analysis will require more sophisticated tools. There are numerous computer programs available to analyze quantitative data. A table of these programs is provided in the appendix. The table provides brief information about the pros and cons of several of the more popular programs.

The implementation team may already have access to one or more of these programs (e.g., Microsoft Excel) or the team may wish to purchase one. Many companies offer free trial versions of their statistical software. These demo versions are typically downloaded directly from the software company's website.

A key factor is having necessary resources to enter and verify the data before commencing analysis. Data from surveys or interviews must be coded, cleaned, and entered before beginning analysis.

Wherever computer programs are used in data collection and analysis (even in simple word processing), the implementation team must *back up* all files on diskettes or CDs, as well as print out hard copies.

Levels of quantitative (or statistical) analysis

The levels of difficulty presented throughout this curriculum are designed to alert the implementation team to the time that may be involved in learning and implementing procedures early in the REA planning process. However, the actual procedures *must be selected because they would best answer the questions that the REA is asking or the information being gathered, not the level of difficulty.*

We will address three levels of quantitative analysis in this chapter. Under the first section - lowest level of quantitative data analysis – we will discuss calculating percents for a "yes/no" (categorical) scale and calculating averages[4] for count-type scales. The medium level of difficulty shown in this chapter will involve comparing the outcomes of two categorical scales in a crosstabulation chart. Finally, the highest level of difficulty will involve performing a statistical test of significance on the crosstabulation introduced in the medium level of difficulty section.

Lowest level of quantitative data analysis

Type 1: Calculating percents. This first level of quantitative data analysis is appropriate for the times when the implementation team has a categorical variable in their survey and they wish to determine the percent of the sample responding in one or more of the categories. Examples of categorical variables are survey questions requiring a “yes/no” response (and sometimes a “don’t know” response), Group (like Intervention versus Comparison), Gender (male and female), and Race/Ethnicity (African American, Asian American, etc.). These categorical variables may have only two categories or many categories, but it is informative to look at the percent of each category to understand the data.

***Preparing data for analysis*.** Each question of the survey is considered a variable on which the respondents provide various responses. These variables are either categorical or numerical. Before applying any calculations on the data, the implementation team should look over the data carefully and “clean” out any obviously bad data. For example, someone may provide several answers to a question where only one is allowed, or an out of range response (such as age=211 years). The process of cleaning the data depends somewhat on what kind of variables are in the survey. Page through the surveys, and scan the responses from each question. If a team member comes across responses that seem dubious, they will need to decide in a systematic manner if the information is correct or not. If it is decided it is incorrect, the implementation team may either try to find the correct response or simply eliminate this respondent’s answer from any future analysis on this question. Sometimes it is clear the response must be incorrect. For example, if a question on your survey asks “In the past year, how many months have you been in stable housing?” and the response is 15, the team must realize this answer is impossible, and it should be eliminated from the analysis. If the implementation team uses computer software for entering and analyzing survey responses, some techniques can be used to detect unusual values more easily. Categorical variables always have fixed-choice responses and are easy to check over. Occasionally, a survey respondent may circle more than one choice for a categorical question’s listed categories. If the categories have some numerical ordering (like on a Likert scale), the average of two circled choices is often regarded as the intended response. This is acceptable as long as the two categories are next to each other. See below for an example of a Likert scale with two circled choices adjacent to each other:

Example: a survey question reads "On a scale of 1 to 5, please rate yourself on the riskiness of your own sexual behavior [CIRCLE ONE]:"

Not at all Risky 1 2 3 4 5 Extremely Risky

We note that the respondent circled two choices next to each other, but rather than lose this response, we assume the respondent meant the riskiness of their sexual behavior is somewhere between 4 and 5, and, thus, we use 4.5 as their response.

If the categories are not next to each other, the typical option is to disregard the response on that survey. See below for an example of the same Likert scale with two circled choices apart from each other:

Example: "On a scale of 1 to 5, please rate yourself on the riskiness of your own sexual behavior [CIRCLE ONE]:"

Not at all Risky 1 2 3 4 5 Extremely Risky

We note that the respondent circled two choices apart from each other. The respondent may have circled one choice, changed their mind and circled another choice – then neglected to cross out the first choice. Or the respondent may feel that their sexual behavior is sometimes somewhat risky and not at all risky other times. Researchers will generally eliminate this response from any data analysis.

Performing the analysis: calculating percents. If the analysis is to be done without the use of a computer, we recommend at least a good, basic hand-calculator be available for use. In addition, have available a notepad for keeping track of the implementation team's calculations. If the implementation team has multiple survey questions to analyze, they will need to keep well-organized notes and calculations. Clearly label each page or section of a page with the question number and variable name from the survey. At least two members of the implementation team should independently perform the calculations to ensure accuracy.

Recall, in the section on Quantitative Data Collection, we presented in the chapter on survey research examples of several questionnaire questions. I repeat here an example, and then show how a simple percentage calculation can be done.

Suppose again, the aim of the study is to find out what a general population in some limited area knows about the transmission of STDs. The first question related to a section of the questionnaire titled Viral Infections is repeated below:

A. *A person can sometimes catch herpes by kissing an infected person.*

1. *True*
2. *False*
3. *Don't know*

To assess the knowledge this sample of participants had regarding this question, it makes sense to merely determine the percent of the participants that circled the correct response (True, in this case). Let us say we had 80 participants complete our questionnaire (and answer this particular question). By leafing through the completed questionnaires, we count, of the 80 questionnaires, 44 circled "True" (and say 24 circled "False" and 12 circled "Don't know"). It is easy to show with a basic hand-calculator that 44 divided by 80 is .55 or 55 percent. Similarly, 24 divided by 80 is 30 percent and 12 divided by 80 is 15 percent. Thus, we could interpret the results by saying "In this sample, 55 percent of the participants knew that herpes can be acquired by kissing an infected person".

Refining the analysis: Display/presentation of percentages. An appropriate display for "yes/no" questions or other categorical scales with more than two categories is a table. To succinctly and clearly display the results of our question about catching herpes by kissing an infected person, we could show the following table:

Catch Herpes by Kissing?	Count	Percent
Yes	44	55%
No	24	30
Don't know	12	15
Total	80	100

Note: it is generally appropriate to round percentages to the nearest whole percent to cut down on information overload to the reader. Occasionally, however, that means whenever there are more than two categories of a variable, the percents may not add up to exactly 100 percent. This is acceptable and understood by most researchers, but be careful: if the percents add up to something more than a couple percentage points different from 100 percent, it means there is a miscalculation somewhere.

If there are few or no missing observations, the table can be simplified even further by eliminating the “Count” column. In this situation it is usually appropriate to have the total number of respondents cited in or near the table. If a reader of the table would like the count of the “Yes” responses, they merely need to multiply 55 percent by the total 80 to get 44 (.55 × 80 = 44). See the chart below for an example display without counts shown for the individual categories.

Catch Herpes by Kissing? (n=80)	Percent
Yes	55%
No	30
Don’t know	15

If the results of several “yes/no” questions need to be displayed, a different form of presentation would be appropriate. Refer to the following table shown below. This table of four “yes/no” questions is from an analysis conducted by Jill Florence Lackey & Associates several years ago. Fifty-four people were surveyed (although several did not answer all the questions). The table only shows the percentage of “yes” responses and does not consider the percentage of “no” and “don’t know” responses.

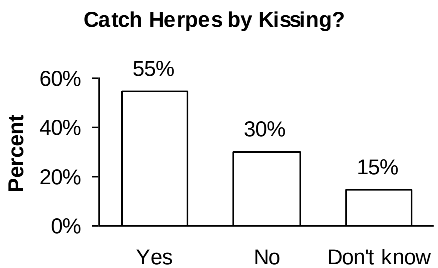
Table 1. Drug use and treatment.

Variable	Percentage "yes"
Are you an IV drug user? (n=54)	98%
Have you entered drug treatment in the past? (n=52)	63
Have you received info about drug treatment? (n=48)	69
Are you in drug treatment now? (n=52)	13

Note: in the table above, the percents do not add up to 100 percent because they come only from "yes" responses to four different questions.

Graphical presentations of the results could also be shown. In general, the benefit of showing a graph instead of a table is that it is more easily digested and understood by people. Downsides are that they take up more space, usually take more time to produce, and often require knowledge of a statistical program. For "yes/no" questions, most researchers will forgo graphs in favor of tables due to the relative ease of understanding percentages.

A graph that might be displayed showing the results of our question about catching herpes by kissing an infected person is a bar graph (shown below).



Interpretation of findings. Generally, the goal of quantitative data analysis is to find evidence from the data that helps answer study questions the implementation team (and others) have developed regarding the nature of their population of interest. Indeed, survey questions are designed to help answer the research questions. If the findings are surprising, the team should first consider the possibility the data were analyzed incorrectly. Occasionally, a scale may have been reversed, and this may lead to a conclusion opposite of what was expected. If the data analysis proves sound, a surprising result may be due to a biased sample – one that is not a good representation of the population. A biased sample is a possibility any time the sample participants are not chosen according to random selection procedures. Finally, a result that does not seem logical may be because the research question was not asked correctly. Perhaps after discussing the surprising results further, logical reasons for these results will emerge. For these reasons, generalizing results from sample data to a population should be done with caution.

The team may now turn to the appendix. Read the brief introduction to Worksheet 3A and complete the practice exercise for this section in which the implementation team will be asked to calculate percentages and produce simple hand-drawn graphs for the provided data.

Type 2: Calculating averages

Performing the analysis: calculating averages. Recall, in the section on Quantitative Data Collection, we presented in the chapter on survey research examples of several questionnaire items. Suppose again, the aim of the study is for the implementation team to develop an STD program that centers on parents talking to their adolescent children about sex and sexually transmitted diseases. The second question related to a section of the questionnaire titled "Sexual experience" is repeated below:

B. How many times did your parents talk to you about sex in the past year?

1. _______ number of times

2. Don’t know

The number put in the blank above would be considered to come from a count scale and would be appropriately summarized with an average. Assume here that 50 adolescents responded to this survey question. After tallying the results of this question the implementation team might summarize them like those shown in the following table:

Number of times parents talked to kids about sex	Count
0	15
1	13
2	7
3	4
4	1
5	2
6	1
7	0
8	1
Don’t know	6
Total	50

Thus, 15 teens responded that their parent(s) did not talk to them about sex or STDs last year, 13 responded their parents talked to them once, and so on. To calculate the average number of times in this sample of teens that parents were reported to talk to their teens about sex and STDs, merely add up the numbers for each response and divide by the number of valid responses. If the tallied results are like those shown in the above table, the sum of the responses could be determined more easily by multiplying each response value by the count (the number of times the response occurred) and adding them up.

For example, (0 x 15) + (1 x 13) + (2 x 7) + (3 x 4) + (4 x 1) + (5 x 2) + (6 x 1) + (7 x 0) + (8 x 1) = 67. Divide this sum by the total number of valid responses – which in this case is 44 (50 – 6 "Don't knows"). The average is 67/44 = 1.523. That is, parents talk to their teens an average of around 1.5 times a year (according to the self-reported data of the teens in this sample).

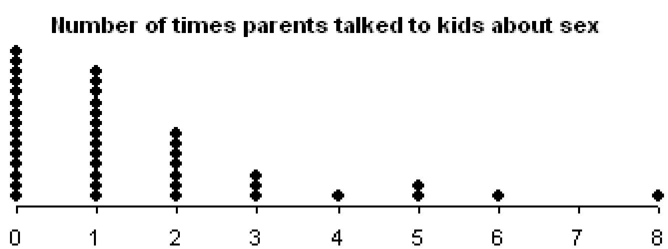
Be aware that extreme numbers may greatly influence the value of the calculated average and may lead to a misleading value of the average. For example, in the example above, if the highest value reported was 48 instead of 8, the average would change from 1.5 to 2.4. As discussed in the section "Preparing data for analysis", unusual responses (called "outliers") should be looked at critically. Does the implementation team believe the datum is accurate? Could the data have been entered incorrectly by the interviewer (or respondent if it was a self-administered survey)? Might the respondent have intentionally put down false information? If a response is highly unusual with respect to the other responses there may be justification in removing that data point from analysis because of the impact it has on the calculations.

Refining the analysis: Display/presentation of averages. If the results of several count-type questions or other questions on a quantitative scale need to be displayed, a form of presentation similar to the one we used for several "yes/no" scales would be appropriate. Refer to the following table shown below. This table of three count-type scales is from an analysis conducted by Jill Florence Lackey & Associates several years ago. Twenty-four people were surveyed. The table shows the average response for the number of times this sample of people reported attending meetings. The total sample of 24 is also split between two groups of people (health consultants and all others)[5].

Table 2: Meetings attended

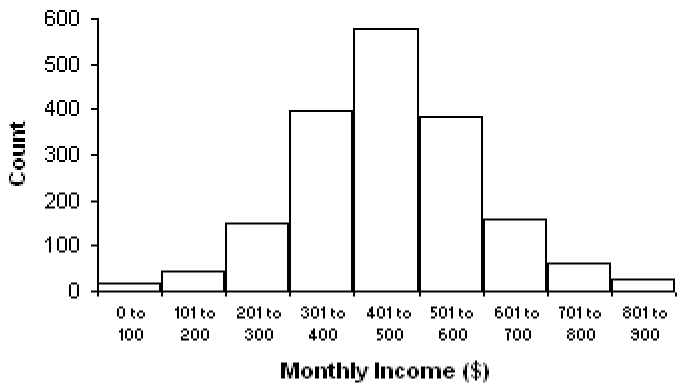
Variable	Average
Number of meetings attended (overall sample) (n=24)	4.6
Health consultants (n=10)	5.4
All others (n=14)	4.0

Several types of graphical displays can be made for quantitative data. If the survey sample is small (say less than 50 people), a simple dotplot would be desirable to show where the data clusters and how they vary. When there is not much data, dotplots are easy to produce by hand. Refer back to the tally of responses for the question about how often parents talk to their teens about sex. To make a dotplot, draw a scale from the lowest to the highest response value of the question. In this example, we would start our line segment at 0 and continue up to 8 in increments of 1 unit. For each response outcome, pile up dots the number of times the response was indicated over the outcome on the line segment. See the example dotplot below. Although this dotplot was produced by a computer, it could easily have been produced by hand within a couple of minutes.



We might interpret what we see in the above dotplot as follows: "The number of times the surveyed teens reported having had their parents talk to them about sex or STDs varies from 0 to 8 times a year, but the majority of the responses are none or only once a year. The data taper off to the right indicating few teens reported having had their parents talk to them multiple times."

For a large sample, a graphic display of a quantitative scale with numerous outcomes could be displayed in the form of a histogram. Histograms combine response outcomes into intervals with the frequency outcome for each interval shown as a bar in a histogram chart. Histograms would ordinarily be produced with statistical software, but can be created by hand with some effort. See below for an example of a histogram showing the response frequencies of reported income for a hypothetical population (or very large sample) of people.



To further expand how we may make use of average calculations, let us return again to the question about how many times parents talked to their teen about sex in the past year. It might be of interest to compare the responses to this question with the responses to the previous question in the survey. That previous question was also originally presented as a sample survey question in the chapter on survey research. It is reprinted below.

A. Have you had one or more sexually transmitted diseases in the past year?

1. No
2. Yes
3. Don't know

The implementation team may hypothesize that teens that responded with a "Yes" to this question would be more likely to have had fewer talks with their parents about sex and STDs. We can see if the data support this "hunch" by calculating the average number of times parents talked to the teens for those that said "Yes" to question A . The team would then compare this to the average number of times parents talked to them for those that said "No" to question A. The team could do this by taking all the surveys in which respondents answered "Yes" to question A[6] and find the average number of times they reported their parents talked to them (from question B). Do the same for the surveys in which the respondents answered "No" to question A. A possible tally and summary of these calculations are shown in the following tables:

Number of times parents talked to about sex	STD in past year? No Count	Yes Count	Total Count
0	7	8	15
1	10	3	13
2	7	0	7
3	3	1	4
4	1	0	1
5	2	0	2
6	1	0	1
7	0	0	0
8	1	0	1
Don't know	2	4	6
Totals	34	16	50

Had at least 1 STD in past year?	Count	Average # of times parents talked to teen about sex
Yes	12	0.5
No	32	1.9
Total	44	1.5

We see the evidence from the data does support our original hunch because, on average, teens that reported not having an STD in the past year reported having their parents talk to them about sex or STDs 1.9 times; whereas, teens that reported having an STD in the past year talked with their parents about sex or STDs only 0.5 times. The implementation team should verify the averages shown in the above table using the method discussed earlier in this section.

A common convention for reporting averages (or other statistics) is to round the calculations to one significant digit more than what the raw data has. The average of 1.9 shown in the above table was actually calculated out to be 1.90625, but we round this number to 1.9, so it is only one significant digit more than our raw data (which consisted of 1 significant digit values ranging from 0 to 8).

***Interpretation of findings*.** Without the benefit of a formal statistical test, we cannot determine for certain if the difference in average number of times parents talked to their teens about sex or STDs calculated above represents a real difference between the groups (teens that have had an STD versus those that have not). Even if the difference is real, interpretation of what it means may be difficult. Statistical testing will be introduced in the context of a crosstabulation example under the highest level of quantitative data analysis section. Although the basic ideas of statistical testing are the same for comparing averages as for comparing counts in a crosstabulation, the technical aspects are more difficult when comparing averages and thus will not be discussed here. We recommend the implementation team use statistical software or consult with an experienced data analyst for conducting statistical tests comparing averages between groups.

The team may now turn to the appendix. If this kind of analysis appears appropriate for the team's study, read the brief introduction to Worksheet 3B and complete the practice exercise. The implementation team will be asked to calculate and compare averages.

The next section will discuss a slightly more complex way of analyzing data.

Medium level of quantitative data analysis

Does the team need to compare data to answer its research questions? Is the team wishing to compare survey data, such as comparing responses by various populations or cultural groups? Does the team need to conduct an experiment where responses from intervention and comparison groups will be compared?

In this section, we analyze the relationship between two categorical scales with a crosstabulation. Recall, in the section on Quantitative Data Collection, in the section on "Experiments," we introduced several types of experiments involving an intervention group and a comparison (or control) group. Many of these experiments also involve a pre-test and a post-test. At this time, we introduced a method for comparing the intervention group with the comparison group for one particular "yes/no" categorical scale only at the pre-test stage of an experiment. This form of analysis may be done prior to an intervention program to see whether the intervention group differs from the comparison/control group on some key attributes.

Preparing data for analysis. The preparation of the data is primarily concerned with cleaning the data and is the same as that discussed in the section, "lowest level of quantitative data analysis."

Performing the analysis: Determining percents of a crosstabulation. Suppose again the aim of the study is to find out what a general population in some limited area knows about the transmission of STDs. The first question related to a section of the questionnaire titled Viral Infections is repeated below:

A. A person can sometimes catch herpes by kissing an infected person.

1. True
2. False
3. Don't know

As described in the section on experiments, we could randomly assign a sample of health educators to one of two groups – the intervention group or the control group. At the pre-test stage of this experiment, both groups should not show any significant differences between each other with regard to their knowledge of STD transmission. We can produce a crosstabulation to see if this is the case.

Assume we have 80 participants able to take part in this experiment. We randomly assign 40 to the intervention and 40 to the control group. After administering the pre-test survey to all 80 participants and tallying the responses for each group we might find, for example, a summary of results that looks like that shown in the table below. This table would be called a 3 × 2 (3 rows by 2 columns) crosstabulation. Only the rows and columns corresponding to the outcomes of the variables are counted when determining the size of the table. The row totals and column totals are called the "marginals"; we will make use of these marginals to calculate percentages associated with this table.

Catch Herpes by Kissing?	Group		Row Totals
	Intervention	Control	
Yes	24	20	44
No	11	13	24
Don't know	5	7	12
Column Totals	40	40	80

The placement of the variables row-wise or column-wise is arbitrary, but Bernard (2006, p. 606) says by convention, the variable that seems to be dependent on the other variable should have its outcomes displayed in the rows.

***Refining the analysis: Display/presentation of data in a crosstabulation*.** To make better sense of the data, percents should be calculated for each of the internal cells of the crosstabulation. Percents can be calculated row-wise, column-wise, or by total. Whether to calculate row percents, column percents, or total percents is somewhat arbitrary, but the choice is made typically based on how one decides to make comparisons. In the table below, we have shown column percents because the intent of the research question is to interpret across the rows (comparing the intervention group to the control group).

Catch Herpes by Kissing?	Group		Row Totals
	Intervention	Control	
Yes	24 (60%)	20 (50%)	44
No	11 (28%)	13 (33%)	24
Don’t know	5 (13%)	7 (18%)	12
Column Totals	40	40	80

The column percents were calculated by taking each internal cell count and dividing it by the column total. For example, for the cell that intersects at the “Yes” row and “Intervention” column, we calculate the column percent of 60 percent by taking the frequency of that cell and dividing by the total number of intervention group respondents. That is, 24 divided by 40 = 0.60 (or 60 percent). Do the same thing for the “Control” column, and we can now compare percents between the two groups. For example, we could say “60 percent of those in the intervention group knew a person could catch herpes by kissing someone infected with herpes, but only 50 percent of those in the control group knew this.” This difference of 10 percentage points seems rather large and provides some evidence that these two groups are already different even before the intervention group received the intervention. We need to decide if this difference is merely due to chance or whether it is due to a real difference in knowledge between the two groups.

We will address this issue with a statistical test in the next “highest level of quantitative data analysis” section, but in the meantime be aware that frequently some minor differences between groups will show up at pre-test stages of an experiment. Generally, these small differences would be attributed to chance, and no fundamental differences between the groups would be assumed – which is important since we want both groups to be roughly the same before one group receives an intervention. We will see in the next section, the results of a formal statistical test on these data will lead us to believe the two groups are not significantly different from each other.

Interpretation of findings. As mentioned in the section under lowest level of quantitative analysis, be cautious when generalizing results to the population from which the sample came.

The team may now turn to the appendix. If the team foresees that comparisons are going to be needed (either in an experiment as this example provided, or for survey data), read the brief introduction to Worksheet 3C and complete the practice exercise. The implementation team will be asked to produce a crosstabulation of counts and percents.

The next section will discuss a more complex way of analyzing data – complex, that is, in terms of this curriculum.

Highest level of quantitative data analysis

Experiments are one of the most sophisticated ways to test an intervention. They are conducted in order to determine whether a treatment has an effect on a participant. A common experiment involves comparing a group of participants that receive an intervention (a treatment - perhaps merely an informational meeting) to another group – the comparison or control group – that will not receive the intervention. Once the data are collected and analyzed, any differences between these two groups must be assessed. The researcher must determine whether the treatment effect is real or not, based upon the results of a statistical test. In this section, we introduce a statistical test that can be applied to a crosstabulation.

Preparing data for analysis. The preparation of the data is the same as that discussed in the section, "Lowest level of data analysis," and again is primarily concerned with making sure the data are ready to be analyzed. This is done when the implementation team makes sure (as much as possible) that no invalid data values get through to the analysis stage.

Performing the analysis: Conducting a statistical test on a crosstabulation. We consider here conducting a formal statistical test on the crosstabulation originally presented in the medium level of difficulty section. The point of this test is to determine whether or not we should believe any relationship we see between the variables of the table is real (statistically speaking). If we do not believe it is real, then we believe the pattern we observed was merely the result of chance. Below, we repeat the 3 x 2 crosstabulation of counts originally presented under the medium difficulty level of data analysis section.

Table of Observed Counts

Catch Herpes by Kissing?	Group		Row Totals
	Intervention	Comparison	
Yes	24	20	44
No	11	13	24
Don't know	5	7	12
Column Totals	40	40	80

We will now call this the table of observed counts – meaning this is what the data actually yielded to us when we tallied the results. Recall, after calculating column percents, we were able to determine at pre-test that the intervention group seemed to be slightly more knowledgeable (60 percent knew the answer was "yes") than the comparison group (50 percent knew the answer was "yes") with regard to this method of herpes transmission. The statistical test now introduced will enable us to decide if that difference of 10 percentage points is real or if it is due to chance.

To assess whether the counts in the table result from a real difference between the intervention and comparison groups, we need to create another table of counts that would be expected to result if there were no difference between the two groups. To get these expected counts, we make use of the following: for each cell of the observed table of counts, multiply that cell's column total by its row total and divide that product by the overall total. For example, to get the expected cell count for the cell that intersects at the "Yes" row and the "Intervention" column, we multiply 44 (row total) by 40 (column total) and divide by 80 (overall total). That is, the expected count for that cell is $\frac{44 \times 40}{80} = 22$. Similarly, we can get the expected cell counts for the other five cells.

The final result is shown below and we will call this our table of expected counts.

Table of Expected Counts

Catch Herpes by Kissing?	Group		Row Totals
	Intervention	Comparison	
Yes	22	22	44
No	12	12	24
Don't know	6	6	12
Column Totals	40	40	80

It makes sense that the farther our observed counts for each cell are away from the expected counts, the stronger is the evidence in support of our notion that the relationship (pattern) between the variables is real. We now introduce a statistic that will summarize the differences between the observed and expected cell frequencies.

We will use the formula $\sum \frac{(O-E)^2}{E}$ where O stands for observed count and E stands for expected count, and we create what is called the chi-square statistic (often denoted as χ^2). The symbol Σ is shorthand notation meaning "sum over". That is, we must sum over all the cells in our crosstabulation taking each O from the table of observed counts and the corresponding E from the table of expected counts. For our example above, we compute the value of chi-square to be

$$\frac{(24-22)^2}{22}+\frac{(20-22)^2}{22}+\frac{(11-12)^2}{12}+\frac{(13-12)^2}{12}+\frac{(5-6)^2}{6}+\frac{(7-6)^2}{6} = \frac{4}{22}+\frac{4}{22}+\frac{1}{12}+\frac{1}{12}+\frac{1}{6}+\frac{1}{6} = \mathbf{0.864}.$$

The bigger this number is, the more likely any difference we see between the groups is real, and not merely chance. But, we need to decide how big is big enough to make this conclusion. Typically, when researchers do a statistical test, they choose to do the test at what's called a 5 percent significance level. This means most researchers are willing to accept a 1 in 20 chance (5 percent) of rejecting the idea that any observed patterns from our data were due to chance, when in fact, the patterns actually were due to chance. Researchers do not know when they make this error, but they do know they make it about 5 percent of the time. We will use a table of chi-square values that were generated from crosstabulations with no relationships to find out whether our data are strong enough to believe the relationship is real (again, in a statistical sense). If the chi-square statistic we computed is at least as large as 95 percent of the possible chi-square values (in other words at least as large as the 5 percent most extreme chi-square values), we should conclude our data are statistically significant.

A chi-square table is provided in the appendix that shows common chi-square critical values. However, before we can use this table one last concept regarding the chi-square analysis needs to be addressed. To look up the correct chi-square value in our table, we must determine our crosstabulation's degrees of freedom. The calculation for the degrees of freedom (df) is straightforward – look at the crosstabulation and merely multiply the number of rows less one by the number of columns less one. A simple formula for this is degrees of freedom (df) = $(r - 1) \times (c - 1)$ where r is the number of rows and c is the number of columns in the crosstabulation (not counting the headings and totals in the margins). In our example, the degrees of freedom is $(3 - 1) \times (2 - 1) = 2 \times 1 = 2$. Thus, our crosstabulation has 2 degrees of freedom.

Now we can look at the chi-square table. Under the row with 2 degrees of freedom and the column that shows 0.050 (corresponding to our chosen 5 percent significance level), we find the chi-square value of 5.991. The chi-square value we computed for our 3 x 2 table is only 0.864, so we should not believe the minor differences we observed between the Intervention and comparison groups are real. We say the results are not statistically significant. Again, in this case, this is desirable because we expect the intervention and control groups to be as similar as possible before the intervention group receives its intervention. After the intervention group does receive an intervention, we expect the data to show differences at the post-test that are statistically significant. Occasionally, some researchers will choose to do their statistical test at a significance level other than 5 percent. We do not discuss the reasons why nor the consequences of doing so here, but the chi-square table included in the appendix contain other columns besides 0.050 so that this can be done.

Interpretation of findings. As mentioned in the section under lowest level of quantitative data analysis, one must be cautious when generalizing results to the population from which the sample came. In addition, any time a statistical test is conducted to compare groups, one must be careful not to attribute causality as the reason for the difference unless the comparison is the result of a randomized experiment (a true experiment). We next discuss the difference in interpretation between a randomized experiment and one that is not a randomized experiment.

Suppose the implementation team would like to compare the average number of past sexual partners between males and females. Although the sample of participants could be randomly chosen from a population, it is impossible to randomly assign participants to be a member of a particular group. Members already have a gender and cannot be forced into the other gender by random assignment. Suppose the data came up statistically significant that the average number of past sexual partners for males is more than the average number for females. Although it is fine to say males, on average, have had more sexual partners than females, we cannot interpret the results to mean being male causes one to have more sexual partners. We reserve the term “causes” only for experiments with statistically significant results in which participants were randomly assigned to the treatments (say intervention and control).

As an example of an experiment, suppose the implementation team wishes to assess the impact of a short STD training course on the knowledge of some healthcare providers. To better judge the effectiveness of the course, an experiment could be conducted in which some providers are randomly assigned to the intervention (the training course) and some to a control group (no training course for them). A statistically significant difference in average score (say from a test on STD knowledge) between the two groups can be attributed to the training course. That is, we could say the training course causes our healthcare providers to be more knowledgeable about STDs. If the healthcare providers were randomly sampled from a population of healthcare providers, we could also generalize the statistically significant results to the population. That is, we could say the training course will help other healthcare providers to be more knowledgeable about STDs.

The team may now turn to the appendix. Read the brief introduction to Worksheet 3D and complete the practice exercise for this section in which the implementation team will be asked to conduct a chi-square statistical test on a crosstabulation.

Getting started

As in data collection, the implementation team can select from three levels of difficulty. However, the level of difficulty should be chosen based on the data analysis strategy that bests suits the team's research design, the purpose of the study, and the data being collected--not on the difficulty factor. The difficulty factor is presented to alert the implementation team to the time and effort involved in these procedures before major decisions are made on study topics and questions.

Lowest level of difficulty**.** Here the implementation team could select one of the analysis strategies summarized under "lowest level of quantitative data analysis," if these strategies fit the study questions. The team should also try and develop at minimum one visual display, such as a table or graph, as this will come in handy when presenting the team's findings. If this option is chosen, begin with the practice exercise in the appendix.

Medium level of difficulty**.** Here the implementation team could select the analysis strategy summarized under "medium level of quantitative analysis," if these strategies fit the study questions. The team should also try and develop at minimum one visual display, such as a table or graph, as these will come in handy when presenting the team's findings. If this option is chosen, begin with the practice exercise in the appendix.

Highest level of difficulty**.** Here the implementation team could select the analysis strategy summarized under "highest level of quantitative data analysis"—again, if this strategy fits the study design and questions. The team should also try and develop at minimum one visual display, such as a table or graph. If this option is chosen, begin with the practice exercise in the appendix.

While chi-square analysis is presented here as the "highest level of difficulty," it should be noted that this is a basic analytic approach covered in introductory statistics training courses. If the team works with trained statisticians, more complex approaches can be used. The addendum at the end of this section also offers a more sophisticated form of quantitative data analysis.

[4]Generally, when people are talking about an “average” they are referring to what is called the arithmetic mean (calculated by summing the data values and dividing by the number of data values). Be aware that some researchers when talking about an average are actually referring to the median – the middle value in the sorted list of data. The median and mean – both measures of a data set’s “typical” value- may be quite different from each other depending on how the data varies. In this course, any mention of “average” will refer to the arithmetic mean.

[5]*Note: displays of tables of averages usually also show in parentheses a statistic called the standard deviation next to the average. The standard deviation is a statistic that summarizes how spread out the data are. A small value of the standard deviation indicates small differences between responses while a large value of the standard deviation indicates wide variation in responses. Because the standard deviation is somewhat difficult to calculate without a sophisticated calculator or computer, it is not discussed as part of the examples shown here.

[6]The “hunch” could be the reverse as well—parents talk to kids about STDs because they are infected.

Section 7, Chapter 3: Quantitative data analysis

7.3.3 Learning activities

Time to review

The implementation team should now perform the following exercises.

1. List two advantages and disadvantages of quantitative analysis.
2. When is the best time to think about using a computer program and/or a trained statistician for quantitative analysis?
3. What are some ways that the team would prepare data for analysis?
4. When would the team be more likely to calculate percents versus averages?
5. What are three ways data analysis can be displayed?
6. When is the best time to use crosstabulation?
7. Why would the team want to create a chi-square statistic?
8. What does it mean to "interpret" findings?

Analyzing quantitative data systematically

Once decisions have been made about the way data will be analyzed, the implementation team should respond to the following questions to check for consistency. See, also, more detailed worksheets in the appendix.

ANALYZING QUANTITATIVE DATA SYSTEMATICALLY

1. Will the team have more than one person involved in data cleaning?
2. Did team members involved in cleaning agree on the data to be cleaned and how they were cleaned?
3. Will the data be entered into a computer for analysis? If so, was the data entry process checked for accuracy by another team member?
4. Does the implementation team have a plan for organizing and storing any data analysis work?
5. Are there key informants in the study that the team can call upon to interpret patterns that the data analysis reveals?
6. Do the data analysts have a start and an end date for the quantitative analysis?

The team is now ready to begin analysis.

Quality control: Checking progress

Once the data analysis is underway, the implementation team should perform quality checks on the analysis at agreed-upon intervals. The researchers can accomplish this by responding to a series of questions. The more detailed worksheets are printed at the end of this chapter.

QUALITY CONTROL ASSESSMENT: QUANTITATIVE DATA ANALYSIS

1. Did the data analysts reach consensus on the cleaning of the data?
2. If the data were analyzed with pencil and paper, have the results been verified by a second team member?
3. If the data were entered into a computer for analysis, have the entered values been checked for accuracy?
4. Is the data analysis output organized and stored appropriately?
5. Has the analysis work been interpreted, and are the interpretations in accordance with the team's expectations?

7.3.4 Resources

Chapter references

Bernard, H.R. (2006). *Research methods in anthropology: Qualitative and quantitative approaches* (4th ed.). Lanham, MD: AltaMira.

Cresswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches* (2nd ed.). Thousand Oaks: CA Sage.

Hardeo, S. and Khurshid, A. (2002). *Pocket Dictionary of Statistics*. McGraw-Hill Irwin.

Pelto, P.J. & Pelto, G.H. (1987). *Anthropological research: The structure of inquiry* (2nd ed.). Cambridge: London.

Additional resources on quantitative data analysis

Dretzke, B.J. (2004). *Statistics with Microsoft Excel* (3rd ed.). Prentice Hall.

Section 7, Chapter 3: Appendix

WORKSHEET CHAPTER 3A: PRACTICE EXERCISE WITH SIMPLE PERCENTAGES

At least two members of the implementation team should be involved in this exercise. The team should expect to arrive at the same answers after completing this exercise.

Assume the team is collecting demographic information from respondents based on the following two survey questions.

A. Gender of interviewee [CIRCLE ONE]
   1. Female
   2. Male

B. Racial/ethnic background of interviewee [CIRCLE ONE]
   1. African American
   2. Asian American
   3. European American
   4. Hispanic/Latino
   5. Native American
   6. Other ____

From a sample of 54 respondents, team members were able to tally the following amounts for the categories of these two variables.

Gender	Count
Female	22
Male	32
Total	54

Race/Ethnicity	Count
African American	13
Asian American	7
European American	16
Hispanic/Latino	10
Native American	3
Other	5
Total	54

1. Working independently, each member of the team should look at the previous tallied information and produce two tables of percents (one for each survey question).

2. Use the information from the tables of percents you have just completed to produce two bar charts showing the percents of each category of the two demographic variables.

See solutions for analyzing this problem below

Solutions to Worksheet 3A

Gender	Count	Percent
Female	22	41%
Male	32	59%
Total	54	100%

or more simply

Gender (n=54)	Percent
Female	41%
Male	59%

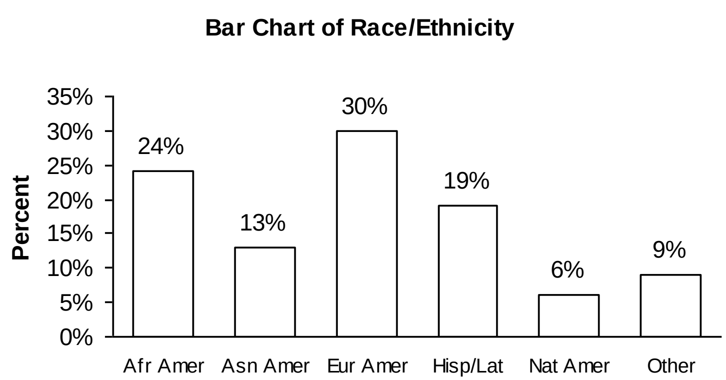
Race/Ethnicity	Count	Percent
African American	13	24%
Asian American	7	13%
European American	16	30%
Hispanic/Latino	10	19%
Native American	3	6%
Other	5	9%
Total	54	101%[1]

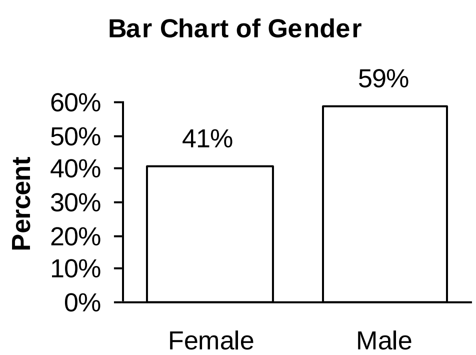
[1] Because of rounding, percentages may not always total 100. They may end up 1 to 2 percentage points above or below 100.

Or more simply

Race/Ethnicity (n=54)	Percent
African American	24%
Asian American	13%
European American	30%
Hispanic/Latino	19%
Native American	6%
Other	9%

Solution to part 2 of Worksheet 3A





WORKSHEET CHAPTER 3B: PRACTICE EXERCISE CALCULATING AND COMPARING AVERAGES

In this exercise, the team will summarize the results of a pre- and posttest survey question. This example comes from the section, "*Data Collection, Quantitative Strategies,*" in the section on "Experiments," in which the goal is to assess whether a housing program for people with HIV has been beneficial health-wise. Below are shown two questions based on Likert scales. The results from a sample of 34 survey participants have been tallied below that.

During the past two months...	**During the same two months one year ago...**
B. ...How consistently did you take your medicine? Never 1 2 3 4 5 Always (as prescribed)	BB. ...How consistently did you take your medicine? Never 1 2 3 4 5 Always (as prescribed)

Response	Count B	Count BB
1	2	4
2	5	5
3	6	12
4	8	6
5	13	7
Total	34	34

1. Determine the average (mean) response ratings for each of these questions. Then, produce a small table for presentation of results.

Solution:

For question B, we could calculate the average like so... $1 \times 2 + 2 \times 5 + 3 \times 6 + 4 \times 8 + 5 \times 13 = 127$. Now, divide this number by the total number of valid responses – which in this case is 34, and the team should get an average equal to $127 \div 34 = 3.735$. Round this to 3.7 for presentation purposes. Similarly, for question BB, calculate the average to be $109 \div 34 = 3.206$. Round this to 3.2.

	Past two months	Same two months one year ago
Take Medicine Rating?	3.7	3.2

or

Consistently take medicine rating	Average
B. Past two months	3.7
BB. Same two months one year ago	3.2

WORKSHEET CHAPTER 3C: PRACTICE EXERCISE
CONSTRUCTING A CROSSTABULATION

In this exercise the team will construct a 2 × 2 crosstabulation in order to compare a sample of males and females about their likelihood of having one or more STDs in the past year. For purposes of this analysis, we will disregard any "Don't knows". Assume the team has 78 surveys with a response to the question on Gender and a "yes" or "no" response as to whether the participant had at least one STD in the past year. Count out that 46 surveys came from males and 32 from females. Also, determine 27 males reported having had at least one STD in the past year, while only 9 of the females reported having had an STD. Construct a 2 × 2 crosstabulation of the counts just provided. Include a column for "Row totals" and a row for "Column totals". After the implementation team has completed this crosstabulation of counts, they should produce appropriate percents for comparing the males to the females. The team may include the calculated percents in the crosstabulation of counts. Finally, in a sentence or two, interpret what the data are saying.

Solution: the first crosstabulation shows counts with row percents

	Had at least 1 STD in past year?		
Gender	Yes	No	Row totals
Male	27 (59%)	19 (41%)	46
Female	9 (28%)	23 (72%)	32
Column totals	36	42	78

or if the team believes “Had an STD” is dependent on Gender (which may not be true, but makes more sense than Gender being dependent on whether they had an STD), the team would organize the crosstabulation like so...

	Gender		
Had at least 1 STD in past year?	Male	Female	Row totals
Yes	27 (59%)	9 (28%)	36
No	19 (41%)	23 (72%)	42
Column totals	46	32	78

This crosstabulation has column percents so that it makes for easy comparison across the rows (that is, between males and females).

Interpretation: “Fifty-nine percent of males reported having had at least one STD in the past year, while only 28 percent of the females reported this.”

Example 2: with numbers easier to analyze

At posttest, the intervention group had received education about STD transmission, but the control group did not.

	Group		
Catch Herpes by Kissing?	Intervention	Control	Row totals
Yes	50 (100%)	30 (60%)	80
No	0 (0%)	20 (40%)	20
Column totals	50	50	100

WORKSHEET CHAPTER 3D: CONDUCTING A STATISTICAL ANALYSIS ON A CROSSTABULATION

The implementation team should go back to the crosstabulation produced for Worksheet Chapter 3C to do this exercise. Conduct a chi-square test on the table to determine whether or not a statistically significant difference exists between males and females with regard to the percent that had at least one STD in the past year.

See solution below.

Table of Observed Outcomes

Had at least 1 STD in past year?	Gender: Male	Female	Row totals
Yes	27	9	36
No	19	23	42
Column totals	46	32	78

Table of Expected Outcomes

Had at least 1 STD in past year?	Gender: Male	Female	Row totals
Yes	21.2	14.8	36
No	24.8	17.2	42
Column totals	46	32	78

The chi-square statistic is calculated to be $x^2 = 7.172$. The degrees of freedom for a 2 × 2 crosstabulation is (2 – 1) × (2 – 1) = 1. Look up in the chi-square table df = 1 and column = 0.050 to find 3.841. Since our result of 7.172 > 3.841 from the table, we say the results are statistically significant. That is, the difference observed between the males and females with respect to having had an STD in the past year is believed to be the result of a real difference between the two sexes. This conclusion could be generalized to the population from which these participants came as long as they were randomly selected.

For our example above, the details of computing the value of chi-square is below:

$$\frac{(27-21.2)^2}{21.2}+\frac{(9-14.8)^2}{14.8}+\frac{(19-24.8)^2}{24.8}+\frac{(23-17.2)^2}{17.2} = 1.587 + 2.273 + 1.356 + 1.956 = \mathbf{7.172}.$$

Example 2 with numbers easier to analyze:

Table of Observed Counts

	Group		
Catch Herpes by Kissing?	Intervention	Comparison	Row totals
Yes	50	30	80
No	0	20	20
Column totals	50	50	100

Table of Expected Counts

	Group		
Catch Herpes by Kissing?	Intervention	Comparison	Row totals
Yes	40	40	80
No	10	10	20
Column totals	50	50	100

The chi-square statistic is calculated to be $x^2 = 25.0$. The degrees of freedom for a 2 × 2 crosstabulation is $(2 – 1) \times (2 – 1) = 1$. Look up in the chi-square table df = 1 and column = 0.050 to find 3.841. Since our result of 25.0 is substantially larger than 3.841 from the table, we say the results are statistically significant. That is, the difference observed between the males and females with respect to having had an STD in the past year is believed to be the result of a real difference between the two sexes.

For our example above, the details of computing the value of chi-square is below:

$$\frac{(50-40)^2}{40}+\frac{(30-40)^2}{40}+\frac{(0-10)^2}{10}+\frac{(20-10)^2}{10} = 2.5 + 2.5 + 10.0 + 10.0 = \mathbf{25.0}.$$

WORKSHEET CHAPTER 3E: ANALYZING QUANTITATIVE DATA SYSTEMATICALLY

1. Will the team have more than one person involved in data cleaning? Yes ______ No _______ It is important to have more than one person involved in data cleaning, so that issues about the quality of the data can be resolved with some consensus among team members.
2. Did team members involved in cleaning agree on the data to be cleaned and how they were cleaned? Yes ______ No _______ If no, bring in another team member or someone more experienced with data analysis and data cleaning to help resolve the problem.
3. Will the data be entered into a computer for analysis? If so, was the data entry checked for accuracy by another team member? Yes ______ No _______ Data entry errors are quite common, so it is important to have someone check over the work of the team member primarily responsible for data entry.

4. Does the implementation team have a plan for organizing and storing any data analysis work? Yes ______ No _______ Data tallied and analyzed by hand should be well-organized in a notebook or binder and kept secured. Data entered into a computer and analyzed with statistical software should be backed up to floppy disks, CD-ROM, or other appropriate storage media and kept secured.
5. Are there key informants in the study that the team can call upon to interpret patterns that the data analysis reveals? Yes ______ No _______ Appropriateness of the analysis and interpretations should be verified by the team members. If the results are difficult to interpret seek the help of an experienced data analyst.
6. Do the data analysts have a start and an end date for the quantitative analysis? Yes ______ No _______ If no, now is the time to set the dates. Start date ______________ Stop date ________________

WORKSHEET CHAPTER 3F: QUALITY CONTROL ASSESSMENT: QUANTITATIVE DATA ANALYSIS

1. Did the data analysts reach consensus on the cleaning of the data?

Yes ______ No _______

If yes, indicate any alterations to the data that took place and why the data were changed.

If no, indicate the problem data points below, and set a future time to revisit this issue to resolve it.

2. If the data was analyzed with pencil and paper, have the results been verified by a second team member? Yes ______ No _______ If no, have someone do the verification because it is easy for one person analyzing data to make a mistake and not discover it. A fresh look at data with a new pair of eyes can avoid an incorrect analysis being carried to the next stage of interpretation.
3. If the data was entered into a computer for analysis, have the entered values been checked for accuracy? Yes ______ No _______ If no, find someone to verify the accuracy of the data entered. Set a time for when this verification is to be complete.

4. Is the data analysis output organized and stored appropriately? Yes ______ No _______ If no, plan to do this. Photocopies of analysis hand-written onto notebook pages would be an appropriate way to make a back-up of the work. Computer storage media should be used for data entered and analyzed with computer software. Keep sensitive information locked up and accessible only to team members.
5. Has the analysis work been interpreted, and are the interpretations in accordance with the team's expectations? Yes ______ No _______ If no, the team should make appropriate interpretations. Keep in mind that how the respondents were sampled may limit generalizations of the results to the population. If the analysis results are not in accordance with the team's research questions, it may be due to an invalid research expectation. Reassess the research questions and their associated expectations.

Chi-Square Table of Critical Values

	Area to the Right					
df	**0.200**	**0.100**	**0.050**	**0.025**	**0.010**	**0.005**
1	1.642	2.706	3.841	5.024	6.635	7.879
2	3.219	4.605	5.991	7.378	9.210	10.597
3	4.642	6.251	7.815	9.348	11.345	12.838
4	5.989	7.779	9.488	11.143	13.277	14.860
5	7.289	9.236	11.070	12.832	15.086	16.750
6	8.558	10.645	12.592	14.449	16.812	18.548
7	9.803	12.017	14.067	16.013	18.475	20.278
8	11.030	13.362	15.507	17.535	20.090	21.955
9	12.242	14.684	16.919	19.023	21.666	23.589
10	13.442	15.987	18.307	20.483	23.209	25.188
11	14.631	17.275	19.675	21.920	24.725	26.757
12	15.812	18.549	21.026	23.337	26.217	28.300
13	16.985	19.812	22.362	24.736	27.688	29.819
14	18.151	21.064	23.685	26.119	29.141	31.319
15	19.311	22.307	24.996	27.488	30.578	32.801
16	20.465	23.542	26.296	28.845	32.000	34.267
17	21.615	24.769	27.587	30.191	33.409	35.718
18	22.760	25.989	28.869	31.526	34.805	37.156
19	23.900	27.204	30.144	32.852	36.191	38.582
20	25.038	28.412	31.410	34.170	37.566	39.997
21	26.171	29.615	32.671	35.479	38.932	41.401
22	27.301	30.813	33.924	36.781	40.289	42.796
23	28.429	32.007	35.172	38.076	41.638	44.181
24	29.553	33.196	36.415	39.364	42.980	45.558
25	30.675	34.382	37.652	40.646	44.314	46.928

Directions for use follow

To use the table for conducting a statistical test on a crosstabulation, determine the desired significance level and the degrees of freedom associated with the table. Look at the intersection of the corresponding degrees of freedom row and significance level column to find the smallest chi-square value needed for the results to be statistically significant. For example, if one has a 4 × 5 crosstabulation and chooses to conduct the statistical test at the 5 percent significance level, go to the row with (4 – 1) × (5 – 1) = 12 degrees of freedom and the column that leaves 0.050 (significance level) area to the right to find the critical chi-square value of 21.026. That is, the chi-square value computed from the actual data from the crosstabulation must be at least 21.026 for the results to be considered statistically significant at the 5 percent level.

Table of computer programs that handle data entry/statistical analysis

Program Name	Web site	Pros	Cons	Other notes
Microsoft Excel	www.microsoft.com	Commonplace; inexpensive	Limited and somewhat difficult to use statistical features	Can be purchased alone or as part of a Microsoft Office package
XLStat	www.xlstat.com	Inexpensive; expands and simplifies Excel's statistical routines	Must have Excel to run; somewhat uncommon outside of academia	An Excel add-in
Analyse-It	www.analyse-it.com	Inexpensive; expands and simplifies Excel's statistical routines	Must have Excel to run; somewhat uncommon	An Excel add-in
Minitab	www.minitab.com	Easy to learn and use; numerous, full-featured statistical routines	Expensive; somewhat uncommon outside of academia	
SPSS	www.spss.com	Somewhat common; numerous, full-featured statistical routines	Expensive; more difficult to learn than Minitab	Most full-featured program; more for serious researchers
Stata	www.stata.com	Somewhat common; numerous, full-featured statistical routines	Expensive; somewhat uncommon outside of academia	

For a list of other statistical software providers go to www.stata.com/links/stat_software.html

Other Resources for Statistical Analysis

<u>Statistics with Microsoft Excel</u> 2/E by Beverly J. Dretzke : Prentice Hall Copyright 2002 ISBN: 0-13-022357-3
This soft cover textbook shows step-by-step how to do basic statistical routines using Microsoft Excel.

<u>TI-83 Plus Graphing Calculator</u> by Texas Instruments
This graphing calculator is extremely popular in academia and has numerous statistical capabilities.

Section 7, Chapter 4: Mixed methods data analysis

7.4.1 Intended learning outcomes

The intended learning outcomes of this chapter on mixed method data analysis follow.

At the end of this chapter, the implementation team will be able to:

1. Understand the advantages and limitations of integrating analysis in mixed methods;
2. Display findings from mixed methods designs in a variety of ways;
3. Integrate analysis of mixed methods in a systematic way.
4. Have a plan for checking the quality of integrating mixed methods at various intervals.

Section 7, Chapter 4: Mixed methods data analysis

7.4.2 Introduction

In mixed methods design, the forms of data analysis will of course be the same as they are within the individual paradigms (quantitative and qualitative). The difference is that the implementation team will be integrating the analyses. Some examples are presented in this chapter.

Advantages and limitations of integrating analysis in mixed methods

***Advantages*.** The overarching advantage of using mixed methods is the opportunity the strategy provides to seek convergence in findings. Convergence is an ultimate sign that the findings are solid. For example, the implementation team might be conducting a study on risk factors associated with a particular STD within a limited size group. The team might be using a concurrent research design where qualitative and quantitative data are being collected at the same time (see section, Data Collection—Mixed Methods). Say that the team collected life histories of a small number of infected individuals who appeared "typical" of the group and also conducted a survey of people in a clinic with the infection. Perhaps they learned through analyzing the life history data that the participants had two experiences in common—recent incarceration and injection drug use. The life histories could provide a detailed process for how the STD was likely to be transmitted, but would not help the implementation team know if this process was widespread. However, if survey results demonstrated that a majority of people infected with the STD had been recently incarcerated and/or used injection drugs, the implementation team has located convergence in the data.

Another advantage of integrating analysis of mixed methods is the richness the strategy yields. In the above example, the researchers can report on the prevalence of the risk factors in the limited-size group, but can also show the processes where these risky behaviors are played out, such as sexual contact and/or injection drug use in correctional facilities.

If the implementation team is using a sequential model, where one method follows another method, other advantages of mixed methods analysis emerge. In the above example, it would be likely that the implementation team was pretty sure of the risk factors associated with the STD (because they had pre-researched the topic), and would then write their survey questionnaire based on that information. However, when little is known about the research question, it is more likely the team would collect and analyze data from either paradigm first before proceeding with the next. For example, suppose the research question was learning the folk beliefs about a certain STD in a recent immigrant population to a neighborhood. The team might begin by conducting focus groups and developing taxonomies on the folk beliefs. These taxonomies could actually be later confirmed through a survey of a wider sample of the immigrant group. In other cases, leading with the quantitative data and analysis might help the implementation team. For example, imagine that the implementation team first wanted to learn the prevalence of misinformation about the STD in the immigrant population. They might conduct a survey, analyze results, and then conduct focus groups to understand why the misinformation exists. One form of analysis guides the next (Creswell, 2003).

A final advantage of integrating mixed methods analysis is the ways it enriches presentation, and convinces audiences. Displays of findings from mixed methods studies can include an almost unlimited inventory of techniques, including tables, quotes, matrices, diagrams, pictures, photographs, and flow charts. Audience members suspicious of “anecdotal evidence” (say from interview excerpts) can look immediately at the numbers. Audience members confused or turned off by statistics can browse through diagrams and quotes.

***Limitations*.** The chief limitation in integrating analysis of mixed methods is an obvious one. The researcher has to know how to analyze both qualitative and quantitative data.

The second limitation is that the process takes longer than analysis of one method alone.

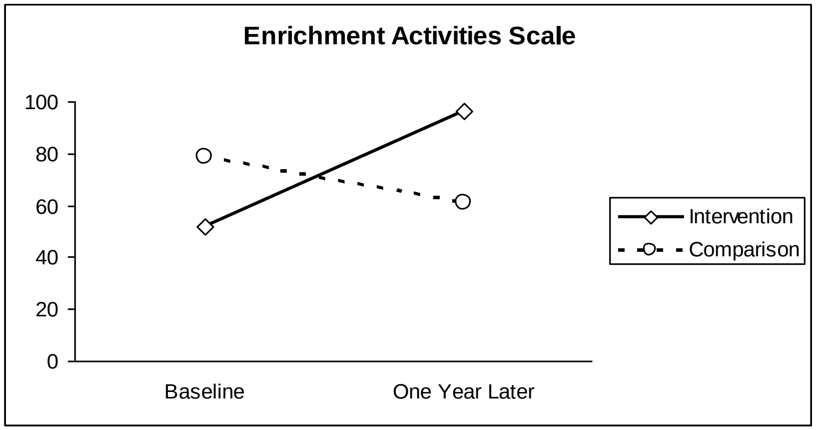
Displaying findings from mixed methods designs

The implementation team should display data from quantitative and qualitative methods in ways that best represent the findings and are easiest for an audience to follow.

Below is an example of an easy way of integrating the methods from an evaluation conducted by Jill Florence Lackey & Associates. The program provided rent subsidies and case management services to clients living with HIV. One of the expectations of the program was that stable housing and rent assistance would give the clients more opportunities to engage in enrichment activities.

The comparison participants, who also have an HIV diagnosis, appear to be generally cutting back in their activities (possibly due to declining health and other barrier behaviors), while the [program] clients are adding new activities. See results below from the pre- and posttest survey.

Average number of enrichment activities over 60 days



During qualitative interviews clients described the ways that program assistance helped them to add enrichment activities.

MADISON: "I saved enough to be able to go out to eat once a week. It is really a treat."

MILWAUKEE: "I decided that now that I have my own place and not a group home I should start a hobby. I started quilting. I have the room and a little money now to buy fabric. It fills up the hours and I can use them for Christmas *gifts."*

MILWAUKEE: "I can do more now with my kids. I don't always feel that good. When I feel good we can go to Chucky Cheese or the movies. Sometimes we rent movies and see them at home."

Another way that analysis of qualitative and quantitative data can be integrated follows. This is one of the matrices shown in Chapter 2 of this section on qualitative data analysis. During this hypothetical study, focus group participants debated the pros and cons of getting AIDS testing. The original matrix is shown first, where only focus group data were collected and analyzed. The second one is a possible result when researchers wished to confirm the results of the qualitative findings with a follow-up survey conducted with a larger random sample of the target community.

Only qualitative findings:

Population trait	**Participants who said they would not get testing**	**Participants who said they would get testing**
Age	Younger (usually under 35)	Older (usually 35 or over)
Social class	Lower (most under middle class)	Higher (most middle class and above)
Sexual orientation	NA (no substantial differences)	NA (no substantial differences)
Gender	Male (usually)	Female (usually)
Neighborhood	Valley Hill, Orangeville	Rigley Homes (mostly)

Integrating qualitative and quantitative findings:

Population trait	Participants who indicated they would not get testing	Participants who indicated they would get testing
Age	Younger (78% under 35)	Older (62% 35 or over)
Social class	Lower (67% under middle class)	Higher (69% middle class and above)
Sexual orientation	NA (no significant differences)	NA (no significant differences)
Gender	Male (62%)	Female (81%)
Neighborhood	Valley Hill, (79%) Orangeville (80%)	Rigley Homes (57%)

The above data displays show ways that researchers can show the convergence of qualitative and quantitative findings. The findings can also be presented in ways that show different dimensions of the same research topic. The possibilities are almost endless. Some examples follow.

- A bar graph showing cases of an STD in several age groups followed by interview excerpts from case managers describing their experiences serving people of these age groups.
- Several paragraphs from observation field notes on the interactions between healthcare workers treating STDs and their patients, followed by a table showing survey results where patients rate the quality of their healthcare.
- A pie chart showing the proportion of different immigrant groups that responded correctly to three questions on the transmission of HIV followed by excerpts from focus groups where each of these immigrant groups discuss HIV.

Section 7, Chapter 4: Mixed methods data analysis

7.4.3 Learning activities

Time to review

The implementation team should now ask each other the following questions.

1. What are two advantages and disadvantages of integrating analysis in mixed methods?
2. What are three ways to display findings from mixed methods designs?

Analyzing mixed methods data systematically

Once the practice exercises for qualitative and quantitative data analysis have been completed, the implementation team should respond to the following questions to check for consistency (also see more detailed worksheets in the appendix).

ANALYZING MIXED METHODS DATA SYSTEMATICALLY

1. Are the team members responsible for analysis now familiar with data analysis from both the qualitative and quantitative paradigms?
2. Will the team have more than one person involved in deciding which analysis displays will be used?
3. Do the data analysts have a start and an end date for mixed methods analysis?

The team is now ready to begin analysis.

Quality control: Checking progress

Once the data analysis is underway, the implementation team should perform quality checks on the analysis at agreed-upon intervals. The researchers can accomplish this by responding to a series of questions. (The more detailed worksheets are printed at the end of this chapter.)

QUALITY CONTROL ASSESSMENT: MIXED METHODS DATA ANALYSIS

1. Have the analysts reached consensus on the most important patterns and findings?
2. Have any key informants been consulted on these key findings?
3. Has the team decided on which data analysis displays will be used?
4. Has the team shown these displays to non-researchers to determine how easy they are to understand?
5. Is the implementation team beginning to discuss the appropriate audiences for these findings?

Section 7, Chapter 4: Mixed methods data analysis

7.4.4 Resources

Chapter references

Creswell, J.W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches* (2nd ed.). Thousand Oaks, CA: Sage.

Additional resources on data analysis for mixed methods

Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches*. Thousand Oaks, CA: Sage.

Section 7, Chapter 4: Appendix

WORKSHEET CHAPTER 4A: ANALYZING MIXED METHODS DATA SYSTEMATICALLY

1. Are the team members responsible for analysis now familiar with data analysis from both the qualitative and quantitative paradigms? Yes _______ No _______ If no, state the plan below for developing this familiarity.
2. Will the team have more than one person involved in deciding which analysis displays will be used? Yes _______ No _______ This collaboration is recommended, but not critical to the study.
3. Do the data analysts have a start and an end date for mixed method analysis? Yes _______ No _______ If no, now is the time to set the dates. Start date ___________________ Stop date _______________________

WORKSHEET CHAPTER 4B: QUALITY CONTROL ASSESSMENT: MIXED METHODS DATA ANALYSIS

1. Have the data analysts reached consensus on the most important patterns and findings?

   Yes ______ No ______

   If yes, briefly describe the patterns below.

   If no, indicate below what the team does agree on (if anything) and set a future time to revisit this issue.

2. Have any key informants been consulted on these key findings? Yes _______ No _______ This is strongly recommended, but it is not always possible, particularly if the team cannot keep names of informants, even in coded forms.
3. Has the team decided on which data analysis displays will be used? Yes _______ No _______ If no, set a date to revisit the issue.
4. Has the team shown these displays to non-researchers to determine how easy they are to follow? Yes _______ No _______ If no, set a date to do this.
5. Is the implementation team beginning to discuss the appropriate audiences for these findings? Yes _______ No _______ If no, now is the time to begin this discussion.

Section 7 Addendum: Advanced Quantitative Data Analysis—Performing Multiple Regression Analysis

Introduction

When conducting the Rapid Ethnographic Assessment, program staff may:

- Want to assess the relative degree to which a number of possible predictive variables influence an outcome of interest in a population.

  For example, in a study of unprotected sexual behavior, the analyst may want to determine whether a number of possible predictive variables are significant, such as sexual orientation, "race"/ethnicity, age group, educational attainment and gender. Findings from such an analysis may help target prevention programming.

- Want to assess the extent to which observed mean differences between sub-groups on an outcome are explained by a third variable.

  Many health status differences attributed to "race"/ethnicity may largely be due to economic differences between the ethnic groups being compared.

- Decide whether a better outcome for a program over the outcome for a comparison group may be confounded by differences between the groups other than the intervention.

  For example, in our initial evaluation of the Delve! curriculum, we examined whether significant differences in post-test knowledge scores between Delve! users and a comparison group (that was provided consultant assistance only) were influenced by differences in the educational level of participants and/or by disrupting intervening events. Multiple regression analysis indicated that the positive outcome for Delve! was independent of these possible confounding variables.

Multiple regression is the analytic strategy of choice for answering questions such as these. It is a general analytic approach, used extensively in quantitative social science research, particularly by economists and sociologists.

How to do multiple regression analysis

Multiple regression is based on the general linear model (this is all the math you will get in this section):

$$y = \alpha + \beta(x) + \varepsilon$$

Where:

y is the dependent (outcome) variable of interest;
α is the intercept of y on the x axis (the point on x where on average y is zero)
β is the slope of y on x; for every unit of x, y on average changes this much; and
ε is the error term or disturbance, the amount that needs to be added or subtracted for the average case to match the actual value of y.

This simple regression model (i.e., a simplified depiction of reality to help us better understand a phenomenon) is expanded on in *multiple* regression. While this sounds very technical, if you think about it carefully and work through the material slowly it is easy to understand conceptually and you will be able to use the results, which is all we really want you to do. The analyst can add several x (independent or predictor) variables, add terms for interactions between predictor variables, or add multiplier terms to account for curvilinear patterns of data. (Curvilinear patterns occur when relationships between measures differ at different points on a distribution, for example a u-shaped curve where there is a strong relationship at the high and low ends of a predictor but not at the mid-range.)

When continuous measures are added to a regression model, there will be several weights (β's) for each case—one for each independent x variable entered into the model. The estimated value for a case is the sum α (constant for all cases) plus the weight for each x times the value of x for that case, plus error (ε). Statistical packages provide a significance test for each of the predictors to assist in determining whether the x variable significantly predicts y or if the relationship observed could be due to chance.

One form of predictor variable (x) is important to discuss. When you have a dichotomy (e.g., yes or no, male or female, person of color or not) coded as either 0 (no) or 1 (yes), the weight for this variable is added to the intercept (α) and essentially changes the point where the line of y on x crosses the zero point on x. Variables such as this are often called indicator or "dummy" variables. Dummy variables can also be included in interaction with continuous variables, such as to look at the relative effect of age on initiation of sexual intercourse among male and female members of a population.

Multiple regression can also handle dichotomous dependent variables by using a variant called "logistic" regression. Logistic regression estimates the odds of an outcome (y) of zero or one for each independent (x) variable in the model. A common dichotomous outcome in HIV/AIDS prevention programming would be whether or not a person admits to engaging in risky sexual behavior. Other forms of distribution of y, such as ordered categories or very rarely occurring events, can also be handled but expert statistical advice should be obtained by non-researchers before doing so. (Actually, any application of regression analysis is likely to need assistance from a statistical analyst for most users of the Delve! curriculum.)

Limitations of multiple regression

Limitations and problems in applying and interpreting multiple regression must also be discussed. While it is often said that regression analysis is "robust" to deviations from its assumptions, there are a number of technical statistical assumptions behind multiple regression that are often violated. This is particularly true in research with small samples or which includes many related predictor (x) variables. While these issues have very technical and statistical explanations, we here provide a basic summary. What is important is that users of regression understand that there are many limitations and nuances to its application and interpretation.

A basic assumption is "no specification error." This means that all relevant variables are included and irrelevant variables are excluded, and that the relationship is in fact linear. Another primary assumption is that the independent variable is not correlated with its error term—that is that there is not a high degree of error at one end of the distribution. However, this may occur when there is poor predictability of an outcome at high levels of a predictor but not at low levels, or the reverse. For example, alcohol consumption generally increases with level of education in the U.S. population—on average, people with higher education are more likely to drink alcohol. However, in a small sample of individuals skewed to those at low educational level, there may be a very high variation in how much those at the lower educational level drink and less variation at the higher level. In this case, the error term would be correlated negatively with the independent measure.

With small samples, cases that are extremes ("outliers") also can cause misleading results. The solution is to always carefully examine your raw data and decide whether some cases are so extreme that they may be incorrectly recorded or otherwise in error, in which case they should be fixed or excluded. Outliers may also be indications that your sample is too small and that if a larger sample were drawn more apparently extreme cases would emerge.

The situation called “multi-colinearity” occurs when there are two (or more) highly related x variables in a model. The x variables essentially can cancel one another out and a stable estimate cannot be obtained. A good example of this occurred when modeling the predictors of tobacco consumption in states in the U.S. Both median income and a measure of educational attainment (percent of adults over 25 with college degrees or higher) were used in a model. While both measures had a high (negative) relationship to tobacco consumption when examined alone, the model was un-interpretable when both median income and education were included, since they were both highly related to each other and essentially cancelled each other out. (In this case, the solution was to create a single “latent variable” of socio-economic status—SES—for each state in the U.S., made up of weighted values of education and income.) This SES variable solved the problem of colinearity between education and income and was, as predicted, negatively related to tobacco consumption.

Recommendations

Beyond very simple models with relatively few variables, the user of this curriculum would do well to consult with a statistician or analyst before attempting to use multiple regression. However, it is important to understand the approach as a consumer of quantitative studies.

If you are using a standard statistical package such as SPSS, SAS, or STATA, multiple regression (including logistical regression) is quite accessible. Excel spreadsheets also can be analyzed using simple regression analysis, which is available in the spreadsheet calculation software.

Multiple regression is a very useful tool in statistical analysis and, once basic descriptive statistics are mastered, regression is the next step in the learning curve.

Example of Use of multiple regression in outcome evaluation:

Jill Florence Lackey and Associates conducted a study of a program to improve the science achievement and scientific career aspirations of middle school mainly Latino and African American girls. Girls were assigned to treatment or comparison groups, with some erosion of the comparison group into the treatment condition. Girls were surveyed annually about their experience in the program, their motivation to continue in science, their knowledge of science, their attitudes about the importance of science, and their intents to pursue a career in science.

In the initial analysis of one-year follow-up data, there were raw mean differences between groups that showed the intervention girls were significantly higher on four of six outcomes than were the comparison girls. This is indicated in

Table A-1 by asterisks in the means column. However, both groups also showed significant before-after change on several of the outcomes, as indicated in the baseline to 1 year p-value column. There was concern that there was differential dropout from the research in the two groups which made them non-equivalent at one year follow-up. There was also some indication that the demographic characteristics of the final groups differed.

Thus a multiple regression analysis was undertaken. For each outcome (y) variable, a multiple regression equation was estimated in which the independent (x) variables were the baseline version of the outcome measure, ethnicity of the students (African American or Latino), the school attending from which assignment to conditions was made, and an indicator variable for condition (0 = comparison, 1= intervention group).

The results of this analysis definitively support the overall benefit of the program on increased science knowledge, confidence in one's own scientific ability, grade point average in science, and career consideration in science. In all of these areas, the one-year score for the intervention students was higher than that for comparison students, controlling for baseline differences, ethnicity and school. Interestingly, this result was positive for the program even on an outcome on which the girls overall declined over time—career consideration in science.

The final column of Table A-1 shows the result of the regression analysis for the intervention indicator variable, in the original metric* of the dependent measure. A variable is a significant predictor at the $p < .05$ level if its coefficient is roughly twice the standard error for the coefficient. Thus in Table A-1 on the *science knowledge* row, we see that the average adjusted difference between the intervention and comparison samples at one year is 1.66 points, on a scale with a mean of about 6 and standard deviation (s.d.) of about 2.3 at baseline. Similarly, science GPA (theoretical range from 0 to 4.0; mean in this sample at baseline of about 3.0 and s.d. of about .76) showed an average adjusted difference at one year of 0.64 points in favor of the intervention students. While career consideration in science decreased in both groups (last row), the intervention group still had a significantly higher mean score.

*NOTE—"standardized" regression coefficients can also be obtained from most statistical packages, which express the results in standard deviation units which can be compared between variables measured in different increments or between different samples.

Table A-1: Example of Regression Analysis for Assessing Program Outcome:
Evaluation of the *Science Explorations* Program

Change in	**Baseline**				**1 year follow-up**						**Multi-variate analysis**
	Intervention[a] (n=132)		Control[a] (n=84)		Intervention (n=132)			Control (n=84)			
	Mean	SD	Mean	SD	Mean	SD	Base-line to 1 yr *p*-value	Mean	SD	Base-line to 1 yr *p*-value	Adjusted[b]B, (s.d.) and p value for program effect at 1 yr
Science knowledge	6.22	2.28	6.24	2.36	7.04*	1.92	<.001	5.48	2.55	.029	1.66 (0.36), p < .001
Science importance	2.41	0.39	2.38	0.41	2.43	0.36	.632	2.34	0.40	.364	.052 (.063), p = .41
Outside support for science	2.20	0.62	2.26	0.62	2.28	0.57	.231	2.42	0.49	.024	-.098 (.086), p = .26
Science confidence	2.00	0.42	2.07	0.50	2.13*	0.40	.002	1.84	0.40	.001	.308 (.066), p < .001
Science GPA	2.88*	0.80	3.22	0.71	3.23*	0.18	<.001	2.66	0.86	<.001	.637 (.114), p < .001
Career consideration in science	1.74	0.48	1.80	0.56	1.65*	0.51	.082	1.49	0.42	<.001	.183 (.075), p = .016

a. Cases with valid data used in the analysis. Baselines without one year follow-up were not analyzed; case loss is n=142 originally assigned to intervention and n=68 originally assigned to control condition. In addition, 16 cases originally assigned to the control group were subsequently placed in the intervention.

b. Multivariate analysis adjusted for baseline value of outcome measure, ethnicity and school to estimate 1 year program effects.

* Unadjusted difference between intervention and control group means significant (p< .05) at this time point.

Resources

Lackey, J.F., Borkan, S.S., Torti, V., Welnetz, T., & Moberg, D.P. (2007). The story behind the findings: Yes, the Science Explorations program worked, but why? *Curator 50*(3).