PART I

WHAT TO CONSIDER BEFORE DRAFTING SURVEY QUESTIONS

The first part of this text is devoted to the planning and predrafting stage of the purposeful survey design process described in Chapter 1. As we illustrate in these chapters, developing an effective survey requires a fair amount of planning and preparation. Dedicating time at this stage of the survey design process is key to designing a survey that will yield rich, nuanced, and useful data. We cannot overstate the importance of this work being done *prior* to drafting questions. In fact, the absence of these steps in our survey design experiences was largely what compelled us to write this book. Working with people who developed surveys without careful consideration and attention to planning and predrafting efforts, resulting in data sets that were sorely lacking in useable information, compelled us to spend considerable time and energy detailing this stage of the survey design process, despite the natural tendency to want to jump directly to brainstorming questions the moment we decide a survey is needed.

PURPOSEFUL SURVEY DESIGN

1. Planning and predrafting
   a. Determining and articulating survey purpose
   b. Understanding what surveys can measure
   c. Understanding survey respondents
2. **Developing questions**
   a. **Sourcing questions**
   b. **Crafting question stems and response options**
   c. **Considering demographic and other sensitive questions**
3. Finalizing
   a. Pretesting
   b. Preparing for administration, analysis, and use

Chapter 2, "Planning and Predrafting," covers the important steps of determining and articulating the purpose of a given survey effort, which requires a solid understanding of what can be measured with a survey. This starts with confirming that a survey is indeed the right tool for a given research effort. This chapter sets the researcher up for the survey design process and emphasizes the importance of careful planning, as Sudman and Bradburn (1982) illustrate:

> Too often questionnaire writers are so caught up in the excitement of question writing that they jump rapidly into writing questions before they have adequately formulated the goals of the research and thoroughly understood the research questions. Many questionnaires constructed by inexperienced people look as if the researchers did not know what they were trying to find out until they saw what they had asked. (p. 13)

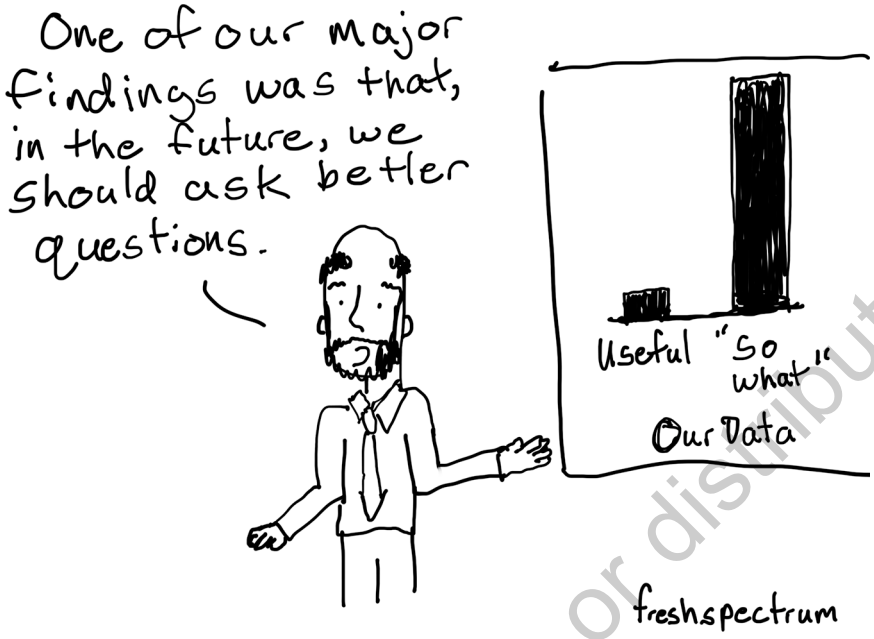
Chapter 3, "Understanding Respondents," provides in-depth coverage of the respondent experience, including challenges they may experience in answering survey questions as well as a discussion of contextual and cultural considerations. This chapter is focused on what survey researchers can do to empathize with and understand respondents prior to developing questions to help ensure a high-quality survey instrument.

2 PLANNING AND PREDRAFTING

WHY SURVEY? THE IMPORTANCE OF ESTABLISHING PURPOSE

Surveys are conducted for a plethora of purposes. A survey may be the only source of information, or it may be part of a multimethod research effort. The stakes may be relatively low, as in a survey that provides information to help an organization understand its clients' interests. Or the stakes may be much higher. Survey results may inform whether a program continues to be funded or is altered in a substantial way (as happens when surveys are used for program evaluations). All surveys should be developed purposefully from the outset, regardless of the centrality of a survey to a given research effort or its potential use for decision making. An unclear or unarticulated purpose is the most common challenge we see in our experience providing consultation to others about survey design. When the purpose of a survey is not particularly clear, it is not surprising that survey questions themselves are unfocused or unclear.

It is wise to begin survey development, or any data collection effort for that matter, by articulating or confirming the key research questions that the data to be collected will help answer. Research questions are the "big-picture" questions that drive and focus a given research effort. Most research efforts aim to answer one or a small handful (often just two or three) of research questions (though more specific subquestions may help identify additional detail desired). It is important to distinguish these questions from questions that are asked in a survey (or in an interview or focus group). It is almost never appropriate to simply pose the research questions to survey respondents. Usually these questions are much too broad to generate useful responses, and, because they are the questions pursued by the researcher, respondents in fact would not be able to answer them. Ensuring that each survey question is in service to these broader research questions helps prevent the capture of unnecessary data and supports the capture of the most useful information possible.



Chris Lysy, Fresh Spectrum

DESIGN DETAILS

RESEARCH AND EVALUATION QUESTIONS

Research questions are high-level questions of inquiry into specific concerns or issues of interest and make up the fundamental core or central element of a study. Research questions focus the study, guide the selection of study methodology, and inform all stages of inquiry, analysis, and reporting. Similarly, evaluation questions focus a program or policy evaluation on the aspects of the program or policy under investigation. Neither of these high-level question types is intended to result in a single data point in the way that an individual survey question would and therefore wouldn't be asked directly to respondents.

Consider the following example: A researcher is studying the use of a new therapeutic technique to determine how well it is supporting improved behavior in patients. The research question might be something like this: "How well is 'Technique X' supporting improved behavior in patients?" A survey of patients could serve as one of several data sources for the research project, but the researcher is unlikely to simply ask patients, "How well is 'Technique X' supporting improved behavior for you?" Instead, the researcher is going to operationalize (or further define) some of the key concepts. For example, what is meant by *supporting*? What is considered

"improved" behavior? What specific behaviors will be explored? Researchers can measure these by asking multiple questions about various ways in which "Technique X" is used and experienced by the patient and about specific ways in which their behavior may (or may not) have changed.

Further, we have found it extremely useful to establish and articulate, in writing, how the survey will help support the broader research purpose and help answer the research questions. In fact, identifying and articulating the purpose for conducting a given survey may be the single most important piece of design advice we can offer. We recommend including the following when articulating survey purpose:

- How the resulting dataset will inform the answer to one or more research questions (i.e., its relationship to the broader research effort)
- Who will use the information captured by the survey
- How the information gathered will be used, including any decisions that rest on the results of the survey
- What form the information will need to be in to maximize its usefulness

> Articulating the purpose for conducting a survey may be the single most important piece of design advice we can offer.

Identifying and articulating the specific purpose for conducting a survey helps focus the process of creating, prioritizing, and selecting questions. As survey questions are generated, revisiting the survey purpose from time to time helps keep the focus of the research *needs* at the forefront, as opposed to addressing *interests* that may not serve the articulated purpose. Pursuing interests that go beyond the survey purpose and research information needs contributes to scope creep, or mission creep, the expansion of a research project beyond its original goals. Scope creep is generally considered harmful to the effort, increasing respondent burden and potentially survey fatigue. For example, researchers may need to understand certain characteristics of people who attend a program, such as where they live, their ages, or why they chose to participate in the program. These may be important given decisions program managers must make about potentially expanding, reducing, or changing the program. It may be of *interest* to know whether the people who attend have a certain level of education, but this may not be *relevant* to the research question(s) or the decisions program managers must be ready to make in this case. Asking questions about education levels *in this case* would only contribute to respondent burden. The point is to clearly delineate *needs* from *interests* based on the research question(s). The survey purpose will also help drive decisions about the format, administration mode (e.g., in person, by mail, online), and length of the survey (more about that in Chapter 8).

WHAT WILL BE MEASURED? UNDERSTANDING CONSTRUCTS AND INDICATORS

Armed with clear research question(s) and an articulated survey purpose, we can start to consider more specifically what constructs our survey can potentially measure. Many researchers make a distinction between "concepts" and "constructs," and we could devote considerable space here to attempting to understand their differences. Some even use one term to define the other. However, in this text, and for our purposes, we use the term *construct* and define it rather broadly as "something to be measured that cannot be directly observed."

Researchers need to define and operationalize constructs in order to measure them in a meaningful way. A piano is not a construct, nor is a cat. These are things that *can* be directly observed. Intelligence, health, prejudice, intent to change, interest in a topic, and awareness or understanding of a social problem are examples of constructs. These are things that *cannot* be directly observed in a person and for which we must identify **indicators**. Indicators tell us the state or level of something. The dots or bars on our smartphones, for example, indicate the level of service we have at any given time or place. Physicians use indicators during a physical examination as evidence of the construct we call "health." Such indicators may include heart rate, blood pressure, cholesterol levels, and body-mass index. These indicators have been identified because health cannot be directly observed. Doctors cannot simply look at us and determine that we are 74% or 92% healthy!

Babbie (1990) refers to constructs, or that which social scientists attempt to measure, as being "rich in meaning," and as such, they "must be reduced to oversimplified, inevitably superficial, empirical indicators. [They have] *no real meanings*, no ultimate definitions" (p. 119; emphasis in original). Consider how people have struggled to define *intelligence*. There is not necessarily a single correct way of measuring constructs like intelligence. It is the researcher's goal to measure constructs in ways that will help answer research questions. Babbie (1990) further contends, "We can never make *accurate* measurements, only *useful* ones" (p. 120; emphasis in original).

> *We can never make accurate measurements, only useful ones.*
>
> *–Babbie*

Let's consider the following example to illustrate these points: A research project aims to understand whether participants in a training session were *satisfied* with the training. *Satisfaction* is a construct. Although a literature search on satisfaction would help a researcher to learn how others have operationalized this construct, it is also possible to identify reasonable indicators that could be measured and provide useful information without reviewing other conceptualizations of satisfaction. The degree to which this is possible depends somewhat on the construct and the context for the research. For example, if the construct to be measured is more complex, like self-esteem or environmentalism, it would be far more valuable (or even necessary) to operationalize (and therefore measure) the constructs in the same way as other researchers, that is, in a research-informed manner.

For our example, how might the construct *satisfaction* be measured in the context of a training session? If participants were satisfied, possible indicators may be that

- they liked the training,
- their time was well spent in the training session,
- they found the materials or resources helpful or useful,
- they felt the instructor was knowledgeable about the topic, and
- they would recommend the training course to a colleague.

Indicators will vary in strength and predictive ability. A dissatisfied participant probably will not report liking the training; hence, liking the training is a stronger indicator. Another participant may agree that the presenter was knowledgeable, but the participant was not satisfied with the training for some other reason (e.g., the content was not relevant). In this case, feeling that the presenter was knowledgeable is a weaker indicator because it does not possess predictive ability. There are certainly many other indicators of satisfaction a researcher could potentially identify.

In gathering information about participant satisfaction, a researcher would likely want to measure other related constructs as well. It may be useful to know whether participants intend to change their practice after a training session, for example. The construct *intent to change* can be operationalized in the same manner as satisfaction. To do so, researchers determine what participants might report if they had an intent to change their practice as a result of the training (i.e., the potential indicators of an intent to change). Once what must be measured is clear, crafting actual questions to measure those constructs (asking about indicators) is much more productive and purposeful.

The following story provides a "real world" example of this stage of survey design. As the researchers in this story refined their questions, they held their survey purpose in mind as they revisited what they wanted to measure. In addition, they empathized with and attempted to understand their respondents to inform the design process and build their prototype.

STORIES FROM THE FIELD

ASKING ABOUT ECONOMIC STATUS IN A SURVEY: AN ODYSSEY

The Institute for Community Health (ICH) is a nonprofit consulting organization that specializes in participatory evaluation, applied research, and strategic planning. We help health care systems, governmental agencies, and community-based organizations improve services and create meaningful impact. We recently encountered a survey design problem that was unexpectedly tricky.

Once upon a time, we set off on a trip to help a client understand *if their program was reaching vulnerable members of low-income communities.* Little did we know, when we started, what a long and winding journey it would be!

At first, we assumed we would define "low income" as "below the federal poverty level (FPL)." We designed a survey that included the following question:

What is your household income? $ _____

To determine whether a person fell above or below the FPL, which varies based on household size, we also had to include:

How many people in your household? _____

However, we found a number of problems with this set of questions.

Household income may be difficult to calculate for some households. Do you know the income of everybody you live with? Do you share all expenses? Do you count the income of your adult child? Do you count the income of someone who only lives there part of the year?

Household size is also a tricky thing to ask for some households! In particular, people may not know how to calculate their household size if they have members who do not live there permanently, or with people who contribute to the household's income but do not live there. This is most common among the economically vulnerable households we expected to be identifying.

Poverty and vulnerability are relative to the cost of living. A household at the FPL is better off in rural Alabama than in New York City, for example.

In our experience, many respondents skip survey questions about income. This could be for the reasons above, but also because of taboos about money and financial vulnerability.

Finally, we didn't need this much information!

We thought of a number of tweaks and repairs to help fix these questions, but we eventually realized that we only needed to know whether a respondent was vulnerable due to their low income. This is fundamentally a yes/no question. So we moved on to the next waypoint in our journey.

We next proposed the following proxy measure for low income.

Do you get ANY of the following benefits: SNAP/food stamps, WIC, SSI, SSDI, TANF, housing assistance, Medicaid:

- O Yes
- O No
- O Unsure
- O I prefer not to answer

We reasoned that in order to receive the benefits above, applicants are put through a rigorous vetting process. In looking at the end results, we would be taking advantage of that vetting. Clever, right? We thought so . . . until we piloted the survey with our clients and grantees.

At that stage, we were faced with an unexpected detour. Our client felt stigma around public benefits would prevent people from answering. Further, they felt some of the respondents could be undocumented immigrants who, even when eligible for benefits, are often reluctant to apply.

Despite these indications that this phrasing was not the destination we were looking for, we lingered for a little while here, attempting the following fix:

Just based on your income, are you or would you be eligible for any of the following benefits: Food stamps, cash assistance, housing assistance, or [state-specific Medicaid name]?

- O Yes
- O No
- O Unsure
- O I prefer not to answer

Perhaps the less said about this half question/half monster, the better.

Coming to our senses, we decided to refocus on the *actual* domain we were trying to assess: vulnerability. The following is based on validated measures of social determinants of health.

In the past 12 months, have any of the following been true? (*check all that apply*)

- ○ You worried whether your food would run out before you got money to buy more
- ○ You worried about losing housing or were homeless
- ○ The electric, gas, water, or oil company threatened to shut off services in your home
- ○ None of the above
- ○ I prefer not to answer

This final question both measures the actual construct we needed to know about (vulnerability) and is responsive to the contexts where it is fielded.

At the beginning of our odyssey, we thought we were looking for the answer to a standard demographic question—income—which had an easy answer. However, through the many stops we made along the way, we learned not to take even the "easy" questions for granted. By the time we reached our ultimate destination, we had gained a new appreciation for the importance of (a) staying connected to the underlying question and (b) the process of piloting.

—Carolyn Fisher, Martina Todaro, and Leah Zallman

WAIT, DO YOU REALLY NEED A SURVEY?

As survey enthusiasts, we sometimes wish we could employ a survey for every research project. It is important, however, to recognize that surveys are not always the best data collection tool for a given research effort. The research question(s), information needs, respondent population, and other considerations may mean another data collection approach is more appropriate. Just as in home remodeling, auto repair, or surgery, great care must be taken to ensure that appropriate tools are selected for the job at hand.

Under the right circumstances, surveys can be just the right tool for the job. As we discussed, it is virtually impossible to know if a survey is the right tool without first establishing *why* the research is being conducted and what is expected to result from it (e.g., knowledge creation, program decisions, customer acquisition). In addition, it is vital to ensure that existing data or a different method of data collection effort would not do a better job of meeting the research needs.

DESIGN DETAILS

PUTTING USE FIRST

We cannot allow readers to go any further in this text without a brief "lecture" about the importance of considering respondent burden and ensuring that we collect *only the data needed* for a given research effort. We can think of countless occasions on which we, or researchers we were working with, wanted to capture more information than was needed

to answer the research question(s). It is always tempting to ask an extra survey question or two ("We may as well ask 11 or 12 questions, as long as we're asking 10 questions"). And sometimes researchers experience a sort of "research FOMO" (or fear of missing out) thinking, *If we don't ask* all *the questions we can think of now, we might not have an opportunity to do so later.* Researchers should resist this inclination in order to limit the burden we place on respondents as much as possible but also because it is simply unethical to gather information that isn't needed. This practice can be viewed as highly intrusive in some circumstances and contributes to the already present challenges of survey fatigue and nonresponse. Researchers should aim to capture only the information necessary to answer the stated research or evaluation questions. Of course, this requires being quite certain about what exactly is needed to answer those questions before administering any survey.

Is a New Data Collection Effort Necessary?

New data should not be collected when the desired information is already available elsewhere and analysis of that data (i.e., **secondary data** analysis) is feasible. Exploring existing data prior to conducting any new data collection effort is worthwhile for several interrelated reasons. As previously mentioned, researchers should only collect data that is truly needed, minimizing respondent burden as much as possible. Researchers should also avoid adding to the existing oversurveying (and subsequent survey fatigue) problem.

The capture and review of secondary data can be especially valuable for research efforts when such data exist and are accessible to researchers. Examples of secondary data use include **case record review** (typically, but not always, managed on paper), analysis of data from administrative databases (typically managed electronically), and use of existing datasets such as from the U.S. Census or American Community Survey. There is almost never a need to use a survey to capture data that already exist in other datasets. Of course, use of secondary data comes with its own challenges as well. There may be multiple barriers to access, and sometimes existing datasets don't contain exactly the information desired. Processing and analyzing these data can be tedious if datasets are especially large or complicated or if data entry was problematic (i.e., resulting in datasets with missing data, duplicated data, or data organized in ways that do not make sense to us).

Is Another Data Collection Tool a Better Fit?

Researchers should consider whether another data collection option might be a better fit. Surveys are often misused and overused, so it is important that researchers consider carefully whether other tools or methods could be more appropriate for a given research effort. It may seem ironic for a book about how to design surveys to promote the consideration of other tools, but we cannot overstate how critical it is to ensure that a survey is indeed the *right* choice and not just the easiest or default option. Even professional researchers occasionally forget that surveys are only one method of collecting **primary data** (i.e., data collected firsthand, directly from respondents themselves). Colleagues have approached us many times requesting help with survey design only to discover through conversation with us that in fact a survey was not actually

the best option given the nature of the information desired, potential respondents, resource constraints, cultural considerations, and other relevant circumstances. The following paragraphs describe some of the most common alternative data collection methods (in addition to secondary data analysis), though this discussion is certainly not exhaustive.

Interviews and Focus Groups. Interviews and focus groups share key characteristics with surveys. Typically, all three methods employ protocols consisting of questions and framing language (e.g., introductory information). However, unlike surveys, interviews and focus groups support the capture of detailed qualitative data. Interviews and focus groups are ideal formats for asking open-ended questions and can be much more unstructured, letting the conversation flow between interviewer and interviewee. Interviews and focus groups also allow for the flexible use of probes. Although in some cases it may be useful to include subquestions that act like probes in a survey, these can be tricky to anticipate and deploy and are not usually used to capture qualitative data.

Interviews and focus groups can also allow the researcher to capture data on facial expressions, tone of voice, body language, and other nonverbal cues that can enhance a rich dataset and allow for deeper insight into the phenomenon being studied. Whenever qualitative data are desired or there is need to clarify a particular question's response in the moment (usually a clue that it isn't possible to develop adequate closed-ended response options for that particular question), researchers should consider an interview, focus group, or other qualitative method. That said, these methods come with the distinct disadvantages of being very time-intensive both to administer and to analyze the resulting data. As a result, fewer respondents can typically be reached through these methods than through a survey, lowering sample size and thus reducing potential generalizability of the resulting information.

Observation. Observation is especially useful when the researcher needs to see firsthand a behavior or practice in action and does not want to rely on the authenticity or accuracy of self-reported data. Although observations are best for accurately measuring behavior, they are often logistically impossible for research studies due to a number of factors, not the least of which is concern for privacy. Resource limitations also often prevent observation from being a viable data collection option, especially because multiple observers may be needed to gather the necessary data over a long period of time, observers need proper training in order to achieve a degree of **reliability**, and the fact that some behaviors are infrequent or not easily observed by another (e.g., how could an observer tell if someone is actually *reading* or just *looking* at a magazine?). Observations are therefore more time-consuming, more labor-intensive (e.g., they include the necessity for the observer to travel to and from sites), and costlier than surveys. Sample sizes are typically limited for observation as well, because researcher time is needed for observing behavior, recording information about that behavior, and analyzing that information.

Tests of Ability or Knowledge. Tests of ability or knowledge are considered by some to be a type of survey, because survey questions can measure knowledge when written appropriately. Such tests or assessments of knowledge are sometimes used in pre-post designs for evaluation of professional development (staff development) or training courses. In these cases, change in

responses can determine the degree to which participants have learned the course content. In contrast, performance tests are tasks in which people are expected to demonstrate their knowledge, ability, or proficiency with a given task that is performed for a reviewer or evaluator. Instead of indicating knowledge through an answer to a question, in a performance test the respondent is asked to actually display or use the knowledge.

DESIGN DETAILS

SUMMARY OF DATA COLLECTION ALTERNATIVES

There are several alternatives to surveys, including the most common methods described in this text:

- *Secondary data analysis* methods such as case record review allow researchers to tap into existing information to answer research questions. Although the greatest advantage in these methods is that no new data collection is required, the information available is sometimes limited, or a great deal of restructuring of information is required in order to analyze the data.
- *Interviews and focus groups* allow for the collection of richer, more nuanced, and detailed qualitative data, but typically fewer respondents can be included and analysis can be time-consuming.
- *Observation* is necessary when a researcher needs to see behavior in action and does not want to rely on self-reported data, but observation data can be labor-intensive both to collect and analyze.
- *Tests of knowledge* are actually similar to surveys but intended to measure actual knowledge or ability (versus perceived knowledge). If well-constructed, these can be quite accurate, but context matters—catch research subjects on the wrong day and their memories can fail them.
- *Performance tasks* allow for observation of ability, demonstrated by the research subject. These can result in rich and highly valid information, but they can also be challenging to develop and administer.

It is important to note that the data collection methods described in this text as alternatives to surveys are certainly not the only possible methods for a given research effort. There are many other methods, including emerging methods such as those from the fields of participatory research and evaluation that researchers may want to consider. Typically, the best way to learn about possible alternative methods is to look for and be open to what methods other researchers have used (or not used) for related topics.

In the following two stories, each of us tells about a time when we needed to carefully consider whether a survey was the right tool, before diving in and crafting questions. In Kim's story, a survey turned out to be the best tool for data collection, whereas in Sheila's story, she helped a group determine that a survey was not the right tool after all.

Once researchers have determined that a survey is indeed the best tool for a given data collection effort, it is important to have a solid understanding of what *can* be measured with a survey, along with the distinct advantages and limitations of using surveys.

WHAT CAN BE MEASURED WITH A SURVEY?

Surveys typically capture self-reported information. In other words, surveys ask respondents to answer questions about themselves and their experiences. As such, surveys can capture the following types of information about or from respondents:

- Attributes (e.g., demographic characteristics such as age, ethnicity, or gender)
- Behaviors (e.g., what people do, such as shop, exercise, engage in hobbies)
- Abilities (e.g., knowledge or skills)
- Thoughts (e.g., attitudes, beliefs, feelings, awareness, opinions, or preferences)

STORIES FROM THE FIELD

WHEN A SURVEY WAS THE ONLY FEASIBLE OPTION

I was once tasked with analyzing case record information from child welfare workers in order to describe and assess the use of kinship placements (i.e., placement of a child with a relative or fictive kin) as an alternative to foster care. Though case record review or secondary analysis based on existing database records would have been the ideal data collection methods for this effort, neither was feasible. Existing data systems were in the middle of a massive technical transition that meant data could not be extracted in a usable format. And existing case records were sitting several states away, and not always understandable to those who weren't the case workers themselves.

And so we devised a somewhat unusual survey. Case records were randomly selected and case workers asked to complete a survey for the cases selected in their charge, reporting the particular details needed for the analysis, including details that weren't tracked in existing online systems. One might consider this a case record review rather than a survey—simply a creative use of an online survey tool to capture case record information. But in many ways it fit the definition of a survey and was the best possible tool for the job at hand.

—Kimberly Firth Leonard

STORIES FROM THE FIELD

WHEN A SURVEY WAS NOT THE RIGHT TOOL

There came a knock on my door one day. It was a district office administrator who needed to use a survey for a new committee he was leading and was pointed in my direction by a colleague who knew me as the resident "survey expert." Excited as always to assist in a survey design process, I readily agreed to meet him and his committee co-chair to get started. As I walked into the conference room, the co-chair had already opened up a textbook related to the committee's topic and was pointing at a list of questions asking if I thought they would be "good questions" for their survey. I laughed and asked if they could start by taking a few steps back and helping me to understand the nature of the committee and why they felt they needed a survey. It turned out that in fact, they did not. The committee was newly formed, and the chairs themselves lacked shared understanding of its purpose and goals. After a lengthy conversation clarifying the committee's purpose and goals and identifying the group's information needs in comparison to available data sources, they agreed that a survey would not be necessary or even appropriate. Instead, we identified other data collection strategies such as conducting a focus group or interviewing a few key informants that would better serve their needs. Lesson learned? Don't get ahead of yourself by drafting questions before making sure a survey is the right choice!

—Sheila B. Robinson

Attributes

Most surveys feature one or more questions about attributes. Attributes are inherent characteristics or qualities such as age, race, ethnicity, nationality, education, employment, religion, gender identity, and sexual orientation. Questions that measure these are called demographic questions (see Figure 2.1 for examples). Attribute questions can also capture information beyond the more common demographic questions, about other quantifiable characteristics. For example, how many years' experience respondents have in an occupational field, what departments they work in, or what organizations they belong to are all examples of attributes that could be captured with well-crafted survey questions.

Despite the common compulsion to capture *all* possible information about attributes, especially classic demographic information (e.g., age, gender, race/ethnicity), it is important to limit the number and type of these questions in surveys. This is especially true for surveys that are intended to be anonymous. It can be possible to identify specific respondents from the combination of answers to questions about attributes. It is worth noting too that there are significant complexities associated with asking questions about more sensitive attributes such as race/ethnicity, gender identity, and sexual preference, and it behooves a survey researcher to carefully consider the most appropriate, respectful, and up-to-date ways to ask these questions. This will depend on the population or sample being surveyed. For more detailed information about demographic questions, see Chapter 6.

FIGURE 2.1 ■ Questions That Measure Attributes

What is your age?

- O 20 or younger
- O 21–30
- O 31–40
- O 41–50
- O 51–60
- O 61 or older

What is your educational level?

- O Did not complete high school
- O Earned a high school diploma or GED
- O Earned an associate's or other 2-year degree
- O Earned a bachelor's or other 4-year degree
- O Earned a master's degree
- O Earned a doctorate (PhD, EdD, JD, PsyD) or other terminal degree
- O Earned other educational credential (please identify the credential)

Behavior

Questions about behavior include those that ask how often or how many times respondents regularly engage or have engaged in certain behaviors of interest (see Figure 2.2 for examples). We can ask which behaviors respondents engage in or do not engage in (e.g., exercise) along with when and where the behaviors take place (e.g., mornings, at a gym). Which behaviors are of interest are determined by the research topic and research question(s) and may range from the seemingly mundane to the unusual. Washing one's hair, consuming alcohol, arguing with a loved one, purchasing groceries, or attending cultural events are just a few of the myriad examples of behaviors that could be of interest to a researcher.

Respondents can typically be trusted to report on their own behaviors via a survey so long as questions are written well enough that respondents are able and willing to answer them accurately, and especially when the survey topic is *not* of a sensitive or threatening nature (e.g., sexual behaviors, use of illicit drugs). For example, respondents must understand exactly what the researcher is looking for regarding questions such as "How often do you drink alcohol?" In this example, respondents may wonder if the researcher is looking for how many days per week or month alcohol is consumed, or how many individual drinks are consumed. Given the question "How often do you read a magazine?" respondents may wonder if the researcher is looking for occasions when they have glanced at a magazine, read an article or two, or read a magazine cover to cover. Any of these could reasonably fall within an individual's interpretation of "read

FIGURE 2.2 ■ Questions That Measure Behavior

On average, how many times a week do you exercise? Exercise is defined as sustained physical activity for the purpose of achieving or maintaining a desired level of fitness with a duration of at least 30 minutes.

- O Less often than 1 day per week
- O 1 day per week
- O 2 days per week
- O 3 days per week
- O 4 days per week
- O 5 or more days per week
- O I do not exercise at all

How many times in the last 3 months did you visit an indoor shopping mall?

- O None
- O 1 time
- O 2–3 times
- O 4–5 times
- O 6 or more times

a magazine." Well-designed questions are clear, leave little room for interpretation, and do not put respondents in a position of having to guess the researcher's intentions. (For more on respondent ability and willingness to answer questions see Chapter 3.) Survey designers must also be clear about which dimensions of behavior to measure, for example, frequency, duration, and intensity. Even measuring just the *frequency* of a behavior demands additional design decisions because it can be broken down into more specific categories:

- Relative frequency (e.g., questions that use "vague quantifiers" such as *sometimes, often, rarely* [see Chapter 5 for more on vague quantifiers])
- Grouped absolute frequencies (e.g., questions that offer intervals such as 1–3 times, 4–6 times, 5 or more times)
- Rates (e.g., 3 or more times per month, a few times per year) (Schaeffer & Dykema, 2011).

Abilities

Surveys can also measure respondents' abilities, skills, or knowledge of a particular subject. However, most of the time when surveyed about abilities, skills, or knowledge, respondents are

really providing information about their *perception* of their own abilities, skills, or knowledge (see Figure 2.3 for examples). This is an important distinction; in many cases, we are not truly measuring abilities but rather *perceived abilities*, and as such, these questions are actually measuring a respondent's thoughts.

In many cases, measuring perceived abilities may be sufficient for a given research effort. However, if a true test of abilities or skills is desired, another research method is likely necessary. Surveys *can* test respondents on *knowledge*, if questions are well-crafted enough to elicit this information. Pre- and posttest surveys may measure respondents' knowledge prior to an event and again following that event. In particular, knowledge questions may follow a professional development (staff development), training, or continuing education course as providers seek to understand the effectiveness or impact of course activities on participants' learning. Knowledge questions can also be part of a **needs assessment** (another type of research effort that often employs surveys) to determine if training is necessary.

FIGURE 2.3 ■ Questions That Measure Abilities (Knowledge)

How many states are in the contiguous United States?

- O 50
- O 48
- O 44
- O 40

Which of the following is an example of a root vegetable?

- O Tomato
- O Carrot
- O Zucchini
- O Cucumber

How would you rate your word processing skills?*

- O Excellent
- O Very good
- O Good
- O Fair
- O Poor

* This question measures *perceived* abilities and therefore might more accurately be described as a question that measures thoughts.

* This question measures *perceived* abilities and therefore might more accurately be described as a question that measures thoughts.

Thoughts

Surveys commonly ask respondents about their perspectives on a topic—their attitudes, beliefs, feelings, awareness, opinions, or preferences about a particular matter. We have collapsed these possibilities into what we call thoughts. Examples of questions about respondents' thoughts include political polling questions about feelings regarding elected officials or employer surveys asking for input about new policies in the workplace (see Figure 2.4 for additional examples). Surveys about public awareness of particular issues like health concerns or social problems also contain questions about respondents' thoughts. Awareness may also be further distinguished from knowledge. For example, a respondent may be *aware* of AIDS as a medical condition and public health concern but may have very little actual *knowledge* about AIDS, such as how it is contracted, its symptoms, or treatments.

Although composing questions that ask about respondents' thoughts may seem easy on the surface, there is much to understand about how people form attitudes, how they articulate them with regard to answering questions about them, and the degree to which they can be influenced by question wording and construction. Attitudes can be fleeting and transient and change over time, or be deeply held and enduring. Though quite common, questions about respondents' thoughts may be the most challenging of all to compose. Chapter 3 includes information about respondent ability and willingness to answer questions that is critically important to consider during the process of drafting and testing survey questions about thoughts.

Schaeffer and Dykema (2011) offer an excellent roundup of question types along with relevant examples. They claim that a list or "compendium" of question types and response dimensions would include the following:

FIGURE 2.4 ■ Questions That Measure Thoughts

In general, how would you rate your experience?

- O Excellent
- O Very good
- O Good
- O Fair
- O Poor

I feel supported by my direct supervisor.

- O Strongly agree
- O Agree
- O Neither agree nor disagree
- O Disagree
- O Strongly disagree

- Questions about events and behaviors (i.e., questions that ask about target objects that occur in time, whether the event is external or internal, performed by the respondent or occurring in the respondent's environment)
- Evaluations or judgments (absolute or comparative) about objects (such questions are often used to measure "attitudes," which we consider a theoretical construction rather than a type of question)
- Internal states or feelings (such as worry or happiness)
- Accounts (such as reasons for leaving a job)
- Classifications (such as whether one has a defined benefit retirement plan)
- Other social characteristics (such as age and marital status)
- Descriptions of the environment (such as how many full-time employees are employed at the place you work)

Many classification questions and social characteristic questions are associated with underlying events or behaviors (e.g., one signed up for the retirement plan, got married, gave birth, etc.). (Schaeffer & Dykema, 2011, pp. 924–925)

These authors' list of question types align well with our categories of attributes, behaviors, and thoughts.

WHAT ARE THE ADVANTAGES AND LIMITATIONS OF SURVEYS?

Surveys are the most common data collection tool for good reason. Surveys can be used to collect large amounts of very specific information from large pools of respondents with a relatively small investment of time, money, and other resources as compared to other data collection strategies. However, surveys cannot capture all possible types of data, in all scenarios, from all people. So, how do we know when a survey is the right tool for a given research effort? We must first understand what surveys can measure (as we have just done) and then consider the general advantages of using a survey, know and understand their limitations, and determine if a survey is indeed the *best* tool for our data collection needs.

Advantages

There are many reasons to consider surveys; here are six that particularly resonate with us. Surveys can do the following:

- Capture needed information with relatively limited resources
- Capture a large dataset formatted for ease of analysis
- Be used to generalize about a larger population from a smaller sample of respondents

- Be used to compare responses between subgroups (e.g., different income levels, genders)
- Be used to collect data across multiple sites or over time to aggregate or compare information
- Be administered in multiple modes

Surveys Can Capture Needed Information with Relatively Limited Resources. Surveys typically require less time and other costs than needed to conduct multiple interviews, observations, or focus groups. Especially with an online survey, the main administration cost is for a subscription to the survey tool. There is no need to pay for postage as would be needed for a postal mail survey, or rent a space or pay additional research staff as might be needed for an in-person survey. Surveys also have the potential to gather data from hundreds, thousands, or even tens of thousands of people simultaneously. However, additional survey costs may include hiring a statistician, if the researcher does not have the capability to perform necessary analyses.

Surveys Can Capture a Large Dataset Formatted for Ease of Analysis. Surveys typically generate standardized data amenable to quantification and relatively easy to analyze. This is especially true when using an online survey tool (e.g., SurveyMonkey or SurveyGizmo), which eliminates the need for data entry as well as some preliminary analysis because basic calculations such as response frequencies are often automatically generated. Most online survey tools also allow for easy export of responses into spreadsheets.

Surveys Can be Used to Generalize About a Larger Population from a Smaller Sample of Respondents. A survey with a high-quality sample of a given population (i.e., one that is adequately large and representative of the population being studied) can be used to generalize information about a whole population. Many public health surveys are good examples of surveys using a targeted, representative sample of respondents to monitor the health status of whole populations. For example, the National Health Interview Survey (NHIS) collects data on a broad range of health-related topics from a large sample of U.S. households. Another example, the American Community Survey (ACS), which regularly gathers more detailed information than its cousin, the U.S. Census, is a longer survey sent to a sampled population (rather than the *entire* population as with the U.S. Census). See the NHIS "National Health Interview Survey" webpage for more information on the NHIS, and the "American Community Survey (ACS)" page on the U.S. Census Bureau website for more information on the ACS.

Surveys Can be Used to Compare Responses Between Subgroups (e.g., Different Income Levels, Genders). With adequate sampling and demographic information, it is also possible to understand whether respondents from different populations or with different attributes answer questions differently. In fact, this is one of the most common ways in which survey data are analyzed.

Surveys Can be Used to Collect Data Across Multiple Sites or Over Time to Aggregate or Compare Information. Because of their typically rigid design (e.g., not allowing for probes

or follow-up questions as might happen during less structured interviews), surveys can also easily be repeated across time or locations. This means that results can also be compared across locations, populations, or points in time, so long as surveys are conducted consistently and adequate information about the context of each survey administration (e.g., when collected or about what location/population) is captured.

Surveys Can be Administered in Multiple Modes. Surveys can be administered in person, over the phone, through the mail, online, or even through a combination of administration modes. With a sharp decline in survey response rates in recent years, many researchers have shifted their thinking about survey administration modes to focus on what mode(s) might reach the most respondents (i.e., what the population would most likely have access and respond to) versus the researcher's preference. This thinking is, of course, in line with a respondent-centered approach to the survey design process.

It is important to note that a high-quality survey designed with well-crafted questions, appropriate and sufficient sampling procedures, and effective administration methods is required for these advantages to come to fruition for any given survey effort.

Limitations

Research typically aims to capture factual, objective information. Thus, it is important to consider whether survey respondents are able to be accurate and honest in their responses. This has implications for what can be captured via a survey, especially given that researchers often have little or no ability to clarify or confirm the information gathered, because many surveys do not ask for identifying information (e.g., respondent names) that would allow for such follow-up. There are types of information that surveys are not well designed to measure and a number of circumstances in which a survey is not the best choice. For example, surveys may not be the best option under the following circumstances:

- When exacting information about a target population's behavior or abilities is needed
- When a substantial amount of qualitative data is needed
- When respondents' nonverbal cues (e.g., facial expressions or body language) are important
- When data are needed from populations not able to respond to a survey

Exacting Information About a Target Population's Behavior or Abilities is Needed. It is important to note that although surveys are often used to measure behavior, they cannot *directly* measure behavior. Questions such as "How many times per week do you consume alcohol?" "How often do you exercise?" or "How well can you ballroom dance?" are examples of questions that actually measure thoughts (i.e., respondents' perception, including memory of how often or how well they do these things), not the behaviors themselves. Direct measurement of behavior

requires observation, or a test of performance that allows a subject (respondent) to demonstrate a given behavior while an observer is present.

> It is also much trickier for respondents to report on the behavior of others. Our own behaviors provide a rich set of experiences, including information about what we wanted to do, what we actually did, and how we felt while doing it as well as information on where the event took place and who else was present (Sudman et al., 1996, p. 203).

Asking respondents about others is generally ill-advised because we are often not in a position to accurately assess and report others' behaviors. The problem emerges because "when asked about the behavior of others, [respondents] may draw on their impression of 'what kind of person' the other is, basing the estimate on an implicit theory of personality" (Schwarz & Oyserman, 2001, p. 142).

> Reports about others' behaviors, often referred to as proxy reports, may be sought because the target person is not available for an interview or because the researcher wants to validate a respondent's reports against the perceptions of a familiar other, often another household member. "Researchers sometimes assume that these proxy reports are more accurate than self-reports. However, controlled experimental studies provide little support for this conclusion. . . . In many cases, respondents may not be fully aware of the others' behavior. Moreover, others' behaviors are even more poorly represented in memory, unless they were extreme and memorable" (Schwarz & Oyserman, 2001, p. 144).

A Substantial Amount of Qualitative Data is Needed. If the researcher needs to capture information that is highly **qualitative** in nature—that which cannot be easily boiled down to specific response options, as would be found in closed-ended questions—a more qualitative method for collecting data such as interviews, focus groups, or a less traditional method may be called for. Interviews and focus groups allow the researcher to take the conversation where it naturally goes and respond to participants' initial answers with probes and subquestions that elicit richer, more detailed responses. Although some probes or subquestions can be designed into surveys with the use of **skip logic**, filtering questions, or conditional **branching** (whereby the particular response to a single question directs the respondent to move to a different or additional set of questions), these are preselected based on a specific set of response options offered versus in the moment as the interviewer or focus group moderator hears a response and chooses to follow up and elicit more data.

Another consideration, especially for self-administered surveys whether on paper or online, is that respondents are often not willing (or able) to provide lengthy or detailed written answers in response to multiple open-ended questions. This happens for a variety of reasons, not the least of which is both the time it takes to organize one's thoughts and to articulate them coherently in writing, along with the willingness to expose one's writing ability to researchers. A more detailed discussion of using open-ended questions in surveys can be found in Chapter 4.

Respondents' Nonverbal Cues (e.g., Facial Expressions or Body Language) are Important. When the research question entails understanding *how* people experience attributes, situations, or programs (e.g., winning the lottery, having depression, fighting cancer, being in a drug rehabilitation program), it may be helpful or even necessary to capture facial expressions, tone of voice, laughter, sighing, or crying along with body language such as gesturing and shifting. In these cases, in-person interviews or video-based methods will be necessary. Although it may be possible to capture notes about nonverbal cues as respondents work through questions during a survey administered in person, it is typically best to design a more qualitative method to allow for integrated capture of this key information along with responses to whatever questions the researcher is posing.

Data are Needed from Populations not Able to Respond to a Survey. Consulting existing literature about the specific target population will be critical in determining the most appropriate methods for capturing information from or with these populations. Special considerations are often necessary when studying children or other vulnerable populations, and surveys are not typically best in these scenarios because they are difficult to adapt for populations with specialized needs. Although there is no hard and fast rule about the exact age at which children can be successfully surveyed, some researchers (see, e.g., Borgers et al. [2000]) suggest that children under 4 years of age should not be surveyed at all. It is important to note that anytime a researcher intends to use a survey with children, extensive cognitive pretesting of both questions and instructions is absolutely necessary. Borgers et al. (2000) offer a brief literature review and specific suggestions for surveying children ages 4 to 16. The following story tells of a time a researcher attempted to collect data for a program evaluation through a survey of young children and the lessons learned in doing so.

One final consideration when determining whether a survey is appropriate is whether the research effort requires asking particularly sensitive questions. In these cases, it is important to consider whether respondents are both likely to and able to answer honestly. It may be best to use a more qualitative approach (e.g., interviews) that allow for greater rapport building and the use of probing questions. We discuss sensitive questions in greater detail in Chapter 6 and feature a story about surveying young children in Chapter 7.

DESIGN DETAILS

NINE REFLECTIVE QUESTIONS TO ASK AS YOU DESIGN A SURVEY

In designing surveys, whether for our own projects or for clients, we ask a series of questions we know will help us develop a high-quality instrument capable of capturing the rich, nuanced data we need. Some of these questions are for the planning and predrafting stage and help us resist the temptation to jump straight into question drafting. Others are for the question development stage where we source or craft our own questions.

1. What is the purpose of the survey? Or what are the main goals in conducting this survey?

   Clarity about survey purpose and goals is vital. This helps set us on the right path to develop questions, to refine (and trim) questions and keep the survey as brief as possible, and to communicate clearly to respondents about why they should respond. We advise writing a brief list or short paragraph stating the purpose and goals and using this as a reference throughout the process.

2. What are the big questions the survey will help answer?

   Do we already have one or more broad, overarching research or evaluation questions that the survey is meant to help answer? If not, we recommend identifying information needs and formulating a few key questions that data from this survey will help answer. These broad questions can help in the same way as our survey purpose, that is, to keep us on track and inform decisions throughout the survey development process.

3. Is a survey the right tool given the overarching goals or purpose? Why?

   Let's think critically about this. We too often jump to surveys as our data collection tool of choice, and they may not be the right one! Clarifying the survey's purpose and identifying the big questions we are trying to answer will help determine whether a survey is the right tool.

   Surveys are useful when we need to gather data from a lot of people or to offer a large number of people an opportunity to offer their perspectives on a topic. They also work well when we have very specific questions to ask, and we can anticipate potential responses (so that we can create specific response options).

   Surveys are often not the right choice if the answers to big questions need to be detailed, nuanced, or richly descriptive, for example. Instead, we might realize that a set of interviews or focus groups would be a better approach.

   Feasibility is also important to consider, in terms of both time and budget. Surveys aren't always the fastest option, or the simplest when it comes to analysis, but they can be efficient if there is a clear purpose, a well-designed survey, and a smart communications plan.

4. Who are we trying to reach? Or who are our potential or desired respondents?

   Who should take our survey? Everyone living in the neighborhood? All adults enrolled in the program? Only department leaders? Only people over 50? Only men? The more we can articulate who we want to hear from and why, the easier it will be to plan for communicating with them about the survey (and why they should respond).

5. What do we know about our desired respondents?

   The more we can learn about our desired respondents, the more we can ensure our survey is easy for them to respond to, or even enjoyable (see Chapter 3 for more on knowing your respondents). Centering respondents' perspectives can help us design questions that are clear and meaningful to those we hope will answer them, which can result in better data.

   Exploring existing data available about desired respondents can also reveal information we don't need to collect! Sourcing data elsewhere can save both us and our respondents time and headaches.

6. How can we best encourage response?

   Knowing as much as possible about respondents can help us determine why, or under what conditions, they might be willing and able to answer our survey questions.

This can also help us determine whether more informal or formal communication methods or approaches would work best, or whether we need help encouraging respondents from someone they know and trust. Or whether offering an incentive is an ethical, reasonable, and feasible choice.

7. How does each survey question (or set of questions) serve the survey purpose or goals? How does it align with the big questions?

   Only questions that clearly serve the survey purpose and goals or can be mapped to overarching research questions should be included. If we find we don't need a question, we don't ask it.

8. Is the question (or set of questions) clear and easy to answer?

   Survey questions that are clear and easy to answer are written in plain language, free of jargon or unnecessarily complicated words or phrases. They will make sense to respondents, given what we know about them, their experiences, backgrounds, etc.

   Closed-ended questions will have response options that match the question stem (the language is aligned), are mutually exclusive, and are inclusive (respondents will be able to find an answer choice that matches their experience). Open-ended questions should make it clear how the respondent is to answer—what is being asked of them.

   The bulk of the rest of this text is designed to help you to write clear, high-quality questions. The Checklist for Quality Question Design in Appendix A is also a handy tool for question design.

9. And finally, are we minimizing respondent burden as much as possible?

   Anything we can do to minimize the burden placed on respondents—either in trying to reduce the demand for cognitive effort by using plain language and clear instructions or in making the survey accessible, engaging, and easy to complete—will help the response rate and is simply kinder to respondents. (See Chapter 8 for more on response rates.)

DISCUSSION QUESTIONS

- Why is it valuable to develop a purpose statement for a given survey?
- How is a survey different from other data collection tools? How can researchers be sure that a survey is the best option for a given data collection need?
- What might be advantages and limitations of surveys that are not discussed in this text?
- What are some potentially sensitive questions or topics that might not be obvious to survey researchers?

DESIGN DRILLS

1. Research Scenario, Part 2 (additional parts are found in other chapters): You are tasked with capturing information that will help assess the effectiveness of a new and ongoing program to support low-income senior citizens. The program is based at public libraries

in several small towns and cities across a large region. Your research team has determined that asking participating senior citizens about their behaviors and thoughts related to the program will provide important data for the project. Is a survey appropriate? Explain your answer. If unsure, what additional information would help you make that determination? What other data collection methods might you consider, and how would you determine whether those are more appropriate than a survey?

2. Your survey: Draft as detailed a purpose statement as possible for your ongoing or anticipated survey project, being sure to connect it to any broader research questions.
3. Your survey: Identify one to three constructs you anticipate needing to measure given your research question(s) and survey purpose. Identify potential strong and weak indicators for each.
4. Create your own research scenarios, one for which a survey would be appropriate and one for which a survey would not be the best tool. Write research questions for each scenario. Exchange scenarios with a colleague or classmate and see if you are in agreement with each other about the use of surveys for each scenario.

EXTENDED LEARNING

For an extended example of operationalizing constructs to inform survey question design:

- Johnson, R. L., & Morgan, G. B. (2016). *Survey scales: A guide to development, analysis, and reporting*. Guilford Press.

For more on surveying youth:

- This article contains a literature review and secondary analysis on surveying children:
  - Borgers, N., de Leeuw, E., & Hox, J. (2000). Children as respondents in survey research: Cognitive development and response quality. *Bulletin de Méthodologie Sociologique, 66*, 60–75.
- This article describes the development of a survey instrument for adolescents ages 11–18:
  - Arthur, M. W., Hawkins, J. D., Pollard, J., Catalano, R. F., & Baglioni, A. J., Jr. (2002). Measuring risk and protective factors for substance use, delinquency, and other adolescent problem behaviors: The Communities That Care Youth Survey. *Evaluation Review, 26*(6), 575–601.
- This article describes a survey for children ages 5–10:
  - Foster, T., & Maillardet, V. (2010). Surveying young patients. *Emergency Medicine Journal, 27*(3), 221–223. https://doi.org/10.1136/emj.2008.065615

Recommended texts and blogs for focus groups and interviews:

- Focus groups:

- Krueger, R. A., & Casey, M. A. (2014). *Focus groups: A practical guide for applied research* (5th ed.). Sage.
- Morgan, D. L. (1996). *Focus groups as qualitative research* (2nd ed.). Sage.
- Roller, M. R., Lavrakas, P. J., & ProQuest (Firm). (2015). *Applied qualitative research design: A total quality framework approach*. Guilford Press.

- Interviews:
  - Brinkmann, S., & Kvale, S. (2014). *InterViews: Learning the craft of qualitative research interviewing* (3rd ed.). Sage.
  - Seidman, I. (2013). *Interviewing as qualitative research: A guide for researchers in education and the social sciences* (4th ed.). Teachers College Press.
  - Spradley, J. P. (2016). *The ethnographic interview* (Reissue ed.). Waveland Press.
- Both focus groups and interviews:
  - Roller, M. *Research design review: A discussion of qualitative and quantitative research design issues* [Blog]. https://researchdesignreview.com