1 WHY QUALITY SURVEYS AND QUESTIONS MATTER

WHAT IS A QUALITY SURVEY?

Let's start with a thorough discussion of what we mean by survey. We define a **survey** as an instrument or tool used for data collection that is composed of a series of questions. Others call this a **questionnaire**; we use the terms *survey* and *questionnaire* interchangeably. Surveys include primarily **closed-ended questions** with distinct sets of **response options**, or answer choices. Though many surveys do employ **open-ended questions**, these do not typically make up the majority of questions included. (See Chapter 4 for a detailed discussion of specific question types.)

Surveys are administered to a group of **respondents**, people who answer those questions. Surveys may aim to collect data from a whole **population** or **census** (as in the U.S. Census, a survey administered to every household in the United States) or from a specific subgroup of people, known as a **sample** (as in the American Community Survey, a survey administered only to select households in U.S. communities). Choosing whether to administer a survey to an entire population or to a sample of that population is a critical survey design decision. Sampling in and of itself is a topic with its own literature base and set of theories and is beyond the scope of this text; however, we offer a number of suggested readings on the topic at the end of this chapter.

Surveys are ideally used to capture information not already available through existing data sources. Surveys can measure respondents' attributes (e.g., demographic characteristics), behaviors, abilities (e.g., knowledge and skills), and thoughts (e.g., attitudes, beliefs, feelings, awareness, opinions, or preferences). Chapter 2 includes a more detailed review of each of these measurement areas (as well as the advantages and limitations of surveys more generally). The distinguishing characteristic of a survey is that the questions are consistently administered, typically in person, through the mail, over the phone, or online. This contrasts at least somewhat with more flexible data collection techniques like interviews.

DESIGN DETAILS

INTERVIEWS AND SURVEYS

We distinguish interviews from surveys, though they can share many similarities. Both are characterized by sets of questions, typically asked in the same manner to each respondent. Both can include open- and closed-ended questions. Here are a few characteristics that distinguish surveys from interviews.

- Surveys may be administered in several modes—the more typical paper or online modes, or in person or by phone with a researcher reading questions. Although the latter seems much like an interview, we still consider these surveys, so long as the instrument used—the questionnaire itself—and how it is administered is structured and standardized.
- **Interviews** can range from semistructured to highly unstructured, letting the conversation flow between interviewer and interviewee. For interviews, researchers are also typically prepared to use **probes**. Probes are additional subquestions we ask such as "Can you tell me more?" or "What exactly do you mean by that?" They are designed to elicit additional information from respondents by encouraging them to add to, expand on, or revise their initial answers.
- Interviews are typically composed primarily of open-ended questions.

Just as with survey design, there is an art and a science to developing and conducting interviews. Although the guidance included in this text may well apply to interviews, there is a significant body of literature specifically on interviews for data collection, including *Interviewing as Qualitative Research: A Guide for Researchers in Education and the Social Sciences*, by Irving Seidman (2013).

Surveys are ubiquitous in modern society, used not only by social science researchers and evaluators but also by marketers, businesses, government, and the media, among others. Most of the examples we include in this text could be considered "special purpose surveys"—those that serve a specific research purpose (Fowler, 2014)—though the bulk of the advice included could apply to any type of survey.

What Makes Survey Questions *Quality Questions?*

At the heart of a quality survey are, of course, quality questions. According to Bradburn et al. (2004), "A good question is one that yields a truthful, accurate answer" (p. 325). Quality questions are relevant and engaging to respondents because researchers have taken the time to carefully craft them to reflect respondents' ability and willingness to answer, and their unique cultures and context. This can only come from an effort to empathize with and understand respondents, which requires knowledge of how respondents comprehend, interpret, and draw from their memories the information we seek. Quality questions are designed with an understanding of the anatomy of open- and closed-ended questions, of how to use language to elicit response, and of how to arrange questions into a coherent measurement tool. Stanley Le Baron Payne (1951), author of one of the earliest texts on survey design, opined, "A 'good' question, among other things, is one which does not itself affect the answer" (p. 72).

> *A good question is one that yields a truthful, accurate answer.*
>
> *–Bradburn, Sudman, & Wansink*

What Makes a Survey a *Quality Survey?*

Quality surveys yield rich, nuanced, useful data that help answer one or more research or **evaluation** questions. (See Chapter 2 for more on research and **evaluation questions**.) Through quality questions, quality surveys effectively engage respondents in providing accurate and useful data. But a survey is more than just a simple set of questions. Sudman and colleagues (1996), prominent survey researchers and authors of many foundational texts on developing surveys, offer "a dual conception of the survey" (p. 1) that we find helpful:

> On the one hand, a survey is a social encounter. On the other hand it is a series of cognitive tasks to be performed by respondents. . . . A survey is a special type of conversation . . . [and] is also a voluntary social encounter between strangers. Thus understanding the rules that govern conversations and social encounters in general should help us understand how survey questions are being understood and answered. (p. 2)

We take to heart the notion that a survey is a form of conversation and social exchange and that respondents are rarely, if ever, *compelled* to answer questions for us. This drives us to use a "respondent-centered" design process. Taking respondent needs, interests, and abilities into account allows us to maintain integrity as researchers; we respect and care for our respondents' well-being and avoid burdening them with unnecessary or unduly difficult questions. Surveys must be more than a quick list of everything a researcher would like to know about respondents dashed off in a series of questions. Instead, we think of surveys as conversations built on a foundation of shared language. To frame that conversation such that we are able to gather the needed information, we must thoughtfully employ a *purposeful* survey design process that takes into account an understanding of and empathy for our potential respondents.

WHY DOES QUALITY MATTER?

Despite the increasing availability of copious amounts of existing data and incredible computing ability to mine that data, surveys are often still the best way to capture the specific information needed for a particular research effort. In particular, surveys are a critical and valuable tool in applied research (and thus, most behavioral and market research). Collecting data about human behavior (such as through direct observation) can be quite expensive and problematic. The proliferation of web-based survey design tools (e.g., SurveyMonkey, SurveyGizmo, Google Forms) have made surveys easier and more cost-effective to develop and administer, greatly increasing their use. In fact, surveys have now become the first line of defense in data collection.

Surveys commonly pop up on retailers' websites when we visit online, and survey links are printed on receipts from brick-and-mortar stores (often accompanied by offers of entry into sweepstakes or the promise of coupon codes). Medical offices and other businesses mail surveys to our homes after in-person visits, and survey requests from banks, airlines, hotels, and others are regularly delivered via email, and occasionally still by telephone.

The rapid rise of the survey in recent years, however, has resulted in an increase in **survey fatigue** that "occurs when survey participants become tired of the survey task and the

quality of the data they provide begins to deteriorate" (Ben-Nun, 2008). This phenomenon is well documented, especially on college campuses where "colleges must prove themselves to accreditors and legislators, and, within campuses, departments contend for scarce resources . . . students have come down with survey fatigue, the main symptom of which is **nonresponse**" (Lipka, 2011).

Chris Lysy, Fresh Spectrum

Despite this, surveys remain one of the most facile means by which we can gain access to and insights about others, especially in great number. Researchers must therefore take great care in deploying surveys in order to combat nonresponse and ensure that respondents provide high-quality data. This now requires even greater attention to crafting high-quality questions and survey tools than is typical; researchers must work with a deft hand, using a purposeful design process to maximize useful survey responses.

WHY DO WE NEED A PURPOSEFUL SURVEY DESIGN PROCESS?

We have identified seven key reasons for a purposeful survey design process. These are culled from our review of existing research as well as from our own experiences.

1. Surveys are inherently imperfect tools and are easily fraught with error.
2. Self-reports are problematic.
3. Surveys are often developed without the benefit of formal study or practice.

4. Surveys developed internally are increasingly common and require extra attention.
5. Question wording can dramatically impact responses and data quality.
6. Today's survey researcher must work to ensure cultural relevance.
7. Survey fatigue and nonresponse negatively impact data collection.

Surveys Are Inherently Imperfect Tools and Are Easily Fraught With Error

Though all researchers strive to ensure the tools they use capture reliable and valid information, no survey or question is entirely immune from error. Further, the existing survey literature has clearly established that even seemingly minor changes in question wording can result in large differences in responses (see, e.g., Bradburn et al., 2004; Sudman et al., 1996; Tanur, 1992; Tourangeau et al., 2000).

Experts before us have categorized several kinds of survey error in a variety of ways. Dillman et al. (2014) provide an excellent description of four kinds of potential **survey error** (coverage, sampling, nonresponse, and measurement), whereas Fowler (2014) describes potential error in two broader categories—errors related to who is answering the question and errors related to the answers themselves. We use Fowler's broader categories to think about possible error because most other types of error fit neatly into this simpler typology. Errors related to who is answering questions may also be described elsewhere as errors related to the interviewer, especially for in-person and telephone surveys. The concepts of coverage, sampling, and nonresponse from Dillman et al. (2014) belong in this broader category. Fowler's (2014) description of errors related to the answers themselves is quite similar to Dillman's description of **measurement error**, defined as "the difference between the estimate produced and the true value because respondents gave inaccurate answers to survey questions" (Dillman et al., 2014, p. 3). Answer-related errors include problems like **bias** and respondent reliability. It is this latter type of error we are most interested in combatting through the advice in this text because this is where the order and phrasing of questions and response options become critical. In addition to the Dillman and Fowler texts, other types of survey error are covered in depth in other survey texts (see, e.g., Alreck & Settle, 2004; Czaja & Blair, 2005).

Self-Reports Are Problematic

Surveys are generally (if not always) self-reports of behaviors, attitudes, perceptions, opinions, knowledge, skills, or attributes. Unfortunately, self-reports can be rife with inaccuracies, over- and underestimations, as well as outright fabrications, particularly when questions are poorly worded or configured. Respondents' own perceptions or their sense of others' perceptions of them may get in the way of accurate reporting. Dillman et al. (2014) break down measurement error into two types: response variance and response bias. One type of response bias, called **social desirability bias**, can make respondents reluctant to tell the truth about any behaviors or attitudes that could be perceived by others as negative or unflattering, or they may be more

likely to overestimate responses about positive behaviors and underestimate responses about negative ones (see Chapter 6 for more on social desirability bias).

Response **variance** is typically the result of the fallible human memory. People don't always have the ability to remember things (e.g., events, feelings, opinions, attributes) exactly as they happened. This is true even when people *think* they are remembering things correctly. This phenomenon can result in responses to the same question varying over time. In other words, reliability in responding accurately can be low. People also tend to exaggerate similarities and differences and over- or underestimate occurrences of behaviors. Humans rely on a complex set of implicit biases to make decisions about how they respond to questions and have tremendous difficulty accurately sorting events and behaviors into specific time periods. (For a more detailed discussion of how memory impacts survey responses, see Chapter 5.) However, many of the challenges inherent in survey responses and the resulting datasets can be mitigated with good question design practices.

Surveys Are Often Developed Without the Benefit of Formal Study or Practice

Research (especially of the "applied" variety, such as program evaluation) is increasingly done by those without formal training or preparation in research methods or survey design. In the general quest for data-informed decision making on various aspects of our work in organizations, people in many types of positions (especially organizational leaders, e.g., directors, department heads, principals) are increasingly expected to gather their own data. And surveys are often the first tool grabbed from the toolkit, especially for those not exposed to or experienced in using other methods to capture data.

It can be helpful to think about why this is problematic by using a driving metaphor. The experience of being a passenger does not automatically result in the ability to drive a car well. We need a certain amount of study and practice behind the wheel to achieve proficiency. That said, it doesn't take years of study or a graduate degree to learn how to drive, and the same goes for survey design. Effective surveys are developed by those who have engaged in *some* study and practice (although we readily admit that we do not have the one-size-fits-all prescription for the "right" amount of study and practice). In fact, the quest for well-crafted surveys evolves as new tools and specific issues around question design arise (e.g., the ever-evolving considerations for how to ask sensitive questions, such as those about gender identity and sexual orientation). Thus, it is vital to continue to hone survey design skills. After all, if we learn to drive as teens but then move to a new city and don't drive for several years, we likely need to relearn some aspects of road safety and perhaps catch up with new laws and the latest in auto technology.

Surveys Developed Internally Are Increasingly Common and Require Extra Attention

Survey design work that is done internally—within organizations and usually about the work of those organizations—brings its own set of complications (as well as being subject to the other concerns discussed in this section). Sometimes these efforts are more informal, as with surveys

used as a quick and easy means to collect perception data from people internal to the organization. And sometimes surveys are developed internally that are used for research about people *external* to the organization. These are typically intended to be as robust as surveys developed by more experienced external consultants. However, when surveys are created internally, even with the right technology in the hands of someone with sufficient expertise and practice, soliciting input from outsiders (or better yet, desired respondents) is especially critical. Otherwise, it is easy to get lost in the familiarity of organizational culture and dynamics, use language that makes sense to people within the organization but not to intended respondents, or to misinterpret responses due to internal perceptions.

Question Wording Can Dramatically Impact Responses and Data Quality

It is critically important for any survey designer to understand the role and extent of the cognitive and communicative processes that respondents go through when answering survey questions. The psychology of asking and answering questions was for a long time largely absent from research, evaluation, and survey textbooks, with just a handful of researchers tackling these issues.

In the 1940s and early 1950s, researchers began experimenting with survey wording and finding that changes in wording produced significant changes in survey results (Schuman & Presser, 1977, p. 152). It soon became clear that for survey researchers to obtain the highest-quality data, potential respondents must be able to easily comprehend, process, and interpret survey questions. In fact, "vagueness and ambiguity can lower response quality and increase measurement error in the survey data" (Lenzner, 2012, p. 2). Examples of vague or ambiguous terms include *sometimes, often, many people*, and *significantly* (for more on these types of "**vague quantifiers**," see Chapter 5) as well as deceivingly familiar terms such as *family, farmer*, and *athlete*. Does an unmarried cohabitating couple "count" as a family? Does a gardener selling some produce by the side of the road count as a farmer? Does a casual weekend bowler count as an athlete? Each word can have multiple and varied interpretations. Even a simple-sounding survey question such as *How many times have you read a magazine in the last six months?* contains a key phrase that is widely open to interpretation. What counts as *read a magazine* could range from glancing through a few pages, to skimming or reading just one article, to reading the issue cover to cover, and everything in between.

STORIES FROM THE FIELD

SIMPLE QUESTIONS ARE NOT ALWAYS SIMPLE

Sometimes, a seemingly simple survey question just doesn't work. I was once asked to analyze data from a brief survey of open-ended questions asked of a small number of practitioners across a few departments in my organization. The department director had

quickly crafted the survey in anticipation of an upcoming meeting where she expected the practitioners to be in attendance and knew she had easy access to them. The first question on the survey was "In what areas do you feel you have expertise?" The purpose of the question was not stated, and no additional direction or elucidation was offered. From the answers received, it was clear to me that the question could have been asked for any number of reasons, and I wondered, *Did the director want to build a database of department members with expertise in different areas? Was there simply a need to get a general sense of departmental knowledge or experience? Something else entirely?*

Let's take a quick look at how a simple question like this can be so problematic. The word *expertise* is wide open to interpretation. What did the director mean by *expertise*? After all, each of these practitioners (the survey respondents) had degrees in their field. Some held multiple degrees, and some had decades of experience. Despite this, many survey respondents crossed out the word *expertise* and substituted *experience* or *strength*. I wondered, *Did they not consider themselves experts? Did humility get in the way of their willingness to identify areas of expertise? Were they concerned that by admitting expertise, they might be called on to teach or train others?* Without talking with respondents about how they approached the question, we will never know.

Perhaps the even bigger issue here is about measurement. What was the director trying to measure when she crafted this question? Was she trying to learn (a) whether people have simply had *training or exposure* in certain areas, (b) whether they have a certain level of *confidence* in a certain area, (c) whether they feel they have enough knowledge or skill to *practice* in a certain area, (d) whether they have the ability to *teach or train* others in that area, or (e) something different? Measuring these would require different questions and speaks to the importance of establishing a purpose for each survey question before crafting the question itself.

And so, after analyzing the resulting data from this question, I wondered, *What do we really know now?* The answer unfortunately was . . . not very much.

—Sheila B. Robinson

This focus on question wording became part of a newer paradigm of survey research that emerged in the 1980s and 1990s and changed the focus on survey error from "a statistical model that focused on the *effects* of survey errors . . . [to a] social scientific model that focuses on the *causes* of survey errors" (Tourangeau, 2003, p. 4; emphasis in original). This shift had survey researchers looking to cognitive psychology and related fields to understand how survey respondents decide whether to participate in surveys and how they go about approaching the complex process of answering survey questions. This is referred to as the cognitive aspects of survey methodology (CASM) movement. CASM began when survey researchers and cognitive scientists convened in 1983 for some rich dialogue and idea exchange on how their worlds could collide and result in some collaborative research. Although one might assume that these two disciplines were closely aligned for years (after all, survey researchers understood as early as the 1940s that question wording could impact responses), this was actually a groundbreaking movement.

Tourangeau (2003) describes CASM with one key sentence: "Reporting errors in surveys arise from problems in the underlying cognitive processes through which respondents generate their answers to survey questions" (p. 5). Hence, Chapter 3 is devoted to unpacking these key cognitive processes and their relationship to respondents' answering questions. At the heart of cognition and shared understanding is language. Sudman et al. (1996) remind us:

> The vehicle through which survey questionnaires are delivered is language. Understanding how people comprehend speech and written materials deepens our understanding of the way in which questionnaire design affects people's answers. Language comprehension is based not only on formal structures like syntax, but more importantly, on pragmatic factors that deeply affect meaning. (p. 2)

This makes it quite clear that comprehensibility plays a huge role in survey design and that "survey data quality is reduced if questions are difficult to understand and exceed the processing effort that respondents are willing or able to invest" (Lenzner 2012, p. 17).

Today's Survey Researcher Must Work to Ensure Cultural Relevance

Challenges and opportunities related to **cultural responsiveness**, relevance, and competence are also important to consider when developing surveys. Researchers increasingly need to attend to the ways that our work operates within and can influence cultural contexts. We must take greater responsibility for ensuring that surveys are reflective of and relevant to our intended respondents. Although language considerations are critical, ensuring responsiveness requires much more than just translating a survey into respondents' preferred language. Among other things, researchers must consider the following as they plan and design a survey:

- If the survey needs to be translated into another language, how can we ensure the translation is valid?
- Does the survey reflect the language used by the respondents, including slang, colloquialisms, and regional dialect?
- Are there subjects that are taboo or inappropriate to ask about in the desired respondents' culture? Or ways of asking questions that would not be well received due to cultural norms (i.e., etiquette)?
- What assumptions or expectations do we have as researchers about the way that respondents will engage with the survey? How can we be sure these don't prevent us from capturing, analyzing, and using data in a responsive, respectful manner?

- How could those fluent in the language be engaged in helping to develop the survey? Helping to analyze or interpret results? How can we otherwise respectfully vet our work with those from the respondent community?
- When using or building on existing validated tools, with what populations were the tools validated? What limitations might those tools have for the population we wish to learn about through our research?

Chapter 3 includes further discussion of cultural responsiveness in survey design, leaning on the theoretical frameworks of **culturally responsiveevaluation (CRE)** and the **Equitable Evaluation Framework (™)**. We have also woven cultural considerations regarding survey question development throughout this text.

Survey Fatigue and Nonresponse Negatively Impact Data Collection

There is tremendous competition for respondents who are oversurveyed and burdened by the constant barrage of requests for perceptions and feedback. Surveys must feel brief and be easy to navigate for the respondent in order for the researcher to have any chance of combating nonresponse and achieving an acceptable **response rate**. Every question must count and yield the most useful data possible. The following tweet was part of a conversation in response to a request posted on Twitter for survey respondents and illustrates a reasonable reaction to a survey with too many questions (see Figure 1.1). Researchers are now forced to cull lengthy lists of potential survey questions, using only those so well designed that they will serve to directly inform key study questions.

Further, key research on survey fatigue indicates that respondents feel oversurveyed based on not only the *number* of surveys they are asked to complete but also the *relevance* of the surveys to their lives (Porter et al., 2004, p. 65). It is now incumbent upon survey researchers to consider how they can address this issue of relevancy with potential respondents and minimize respondent reluctance. Not only does this mean ensuring the survey is culturally responsive, it may also require providing specific incentives for response. A commitment to sharing the results or information about how the results will be used in ways that potentially impact those respondents can also go a long way in motivating respondents. For more on respondent willingness and ability to participate in survey research, see Chapter 3.

FIGURE 1.1 ■ Social Media Post About Long Surveys

Source: Seemant Kulleen

DESIGN DETAILS

RESPONSE RATES

Response rates, though important, should be considered only one indicator of the quality and usefulness of survey data. It is commonly believed that the higher the response rate, the more likely the data are representative of the population the researcher intended to learn about or from, and are therefore more accurate. Some studies have provided evidence that lower response rates may not mean less representative data, and even that extraordinary efforts to raise response rates can result in more skewed data (Krosnick [1999] outlines several). Even the venerable Pew Research Center (n.d.-a) states that "fortunately, low response rates are not necessarily an indication of nonresponse bias."

Despite it being more complicated to understand the degree to which survey respondents represent the population intended to be reached through the survey, calculating response rates is still valuable. The American Associates for Public Opinion Research has a handy list (search online for "American Associates for Public Opinion Research Best Practices") and has developed the standard in calculating response rates. An Excel workbook is available on their website to calculate response rates.

Survey researchers must also determine what characteristics of the population are important given the survey topic, and what can be known about both the population and the sampled respondents. Researchers can then strive to collect pertinent information in order to learn as much as is feasible about the representativeness of the respondents for a given survey (such as through use of demographic questions).

The key reason for using a purposeful approach to survey design is that an effective, successful survey is much, much more than a simple set of questions, quickly drafted after a brief brainstorm and delivered to a group of potential respondents. Consider the case made for customized survey designs (those tailored to the circumstances of an individual survey) in *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method* by Dillman et al. (2014):

> Our response to those seeking answers to specific situations . . . is that there is not a simple set of design procedures that if applied to every situation will be most effective in reducing survey error. The populations to be sampled and surveyed, the kinds of questions that need to be asked, the resources available for doing the survey, and other constraints imposed by survey sponsorship differ greatly across individuals and organizations who wish to do surveys. (p. 16)

In other words, there is no single way to determine a sample, no perfect set of survey questions, and no ideal administration mode for all scenarios. Determining what methods and questions to use for a given survey requires a purposeful design process as well as a deep research-based understanding of surveys themselves.

There is no single way to determine a sample, no perfect set of survey questions, and no ideal administration mode for all scenarios.

APPLYING DESIGN THINKING TO SURVEY RESEARCH

We have chosen the term *design* as a helpful heuristic in keeping with our intent to develop surveys with the mindset of a designer. Many modern designers, whether their work is on products or services, think of design as a "human-centered" or "user-oriented" process. In fact, there is an entire field and associated process of "user-experience design" that is rooted in human–computer interaction and devoted to improving user experiences regarding usability, accessibility, and even pleasure and satisfaction. Our position is that a survey design process focused at least as much on our respondents' experience as on our own (or our clients') information needs will result in richer and more useful data. Before delving into our survey design process, we offer a brief introduction to the concept and phases of a design thinking approach and delineate how each phase dovetails with survey design.

The question development process (and larger survey development process) can be greatly enhanced by applying principles of design thinking. **Design thinking** is both a process and a mindset or way of thinking. Design thinking can be broadly applied to problem-solving challenges in a multitude of circumstances; objects of design can be as diverse as a chair, a school curriculum, a hiring process, or a city garden. Design thinking is generally composed of a set of phases during which designers attempt to deeply understand end users and their experiences, especially in the context of the identified problem. Design thinking calls for making sense of insights gleaned from these efforts, brainstorming solutions to the problem at hand, building prototypes, testing designs, and using feedback to make revisions and refinements. The hallmark of design thinking is a human-centered approach in which the designer puts ego and self-interest aside to design a "product" that best meets users' needs. Regarding survey design, we think of this human-centered approach as a "respondent-centered" approach; a respondent-centered mindset encourages us to put our respondents' needs, interests, and contexts ahead of our own.

Elevating respondents' interests in the survey design process may feel counterintuitive to researchers. Researchers are much more accustomed to having our information needs drive such processes. However, survey questions that are poorly understood, misinterpreted, or skipped by respondents pose serious problems. We have both had experiences with surveys that did not work for various reasons and resulted in a lack of useful data. Some survey efforts failed, and others resulted in low response rates. Looking back, we can see that in some cases we clearly had designed poor questions, often because we did not know or understand our respondents well enough.

Design thinking phases generally include these: empathize, understand, brainstorm, prototype, and test. There are variations of these phases in the design literature, and some phases are also known by other names. We describe each next, along with their specific connections to survey design.

DESIGN DETAILS

DESIGN THINKING PHASES

1. Empathize
2. Understand
3. Brainstorm
4. Prototype
5. Test

Empathize

The first phase in design thinking is *empathize*, during which we get to know as much as we can about potential respondents in order to inform question design.

> Empathy is the centerpiece of a human-centered design process. The Empathize mode is the work you do to understand people, within the context of your design challenge. It is your effort to understand the way they do things and why, their physical and emotional needs, how they think about the world, and what is meaningful to them. (Hasso Plattner Institute of Design, n.d.)

Much of what we know about how people generally respond to survey questions is a result of research on survey design and response. Key studies we have reviewed from the CASM era as well as from more recent times include those that explore the following:

- Question wording
- Response options and rating scales
- Cognitive tasks associated with answering questions
- Survey fatigue
- Nonresponse

Comprehensibility is the defining characteristic of a high-quality survey question (as we explore more deeply in Chapter 3). The more we can learn about our respondents, the better we can craft questions they will be able to comprehend and answer. We suggest gathering as much detail as possible about the specific respondent pool *before* drafting survey questions. What are their demographic characteristics? Knowing about their age, gender, occupation, race/ethnicity, education level, literacy level, income, access to technology, interest in the topic, or even willingness to respond to a survey, along with other attributes, can significantly inform the design process. This can mean the difference between low-quality data and a highly successful survey research effort. Here are some additional questions to consider:

- What are the literacy levels of respondents?
- How are their days structured (at work or at home)?
  - Will they feel they have the time to complete a survey?
  - When are they most likely to be able and willing to complete a survey?
  - How much time might they be willing to devote to completing a survey?
- What are their interests? Do those interests relate to the survey topic?
- What do they value? Do those values relate to the survey topic?
- How much do they care about the survey topic? What aspects of the topic might they most care about, given the context (e.g., home, work, community) in which they will be surveyed?
  - Will they be interested in providing data?
  - Are there reasons they may want or not want to participate in a survey?
  - Will they be interested in the results of the study?
- What is their relationship to the researchers, if any? What is their relationship to the organization conducting the survey (i.e., are they employees, customers, clients, fellow community members)?
  - Do they know us as internal or external to the organization?

Designers typically use in-depth **observation** and interview strategies to learn more about their end users. Survey researchers may not be afforded the luxury of time or resources to empathize with potential respondents as deeply as we would like. However, we are often able to read about the target population as a result of previous research efforts or to ask questions of them, or of those who interact with them more regularly. In survey research, the empathy phase may be fairly straightforward, especially if we know our respondents well or share key characteristics with potential respondents. For example, this may be the case when conducting surveys from inside an organization to be used inside the organization. However, even when it seems we are close enough to the respondents to understand them adequately, researchers must take care not to make assumptions or jump to conclusions about respondents. In particular, we need to know about respondents *in the context of the specific survey topic*.

DESIGN DETAILS

EMPATHY MAPS

In doing the research for this text, we stumbled on a wonderful idea that can easily be adapted for survey design. In a design blog post, "Interaction Designer" Matthew Weprin (2016) introduces "empathy maps" as a way to gather and organize information about end users (e.g., survey respondents). In doing so, he lays out the steps for building the map using chart paper and sticky notes.

How to Build an Empathy Map

1. Draw the map and its four quadrants: Says, Thinks, Feels, and Does.
2. Sketch your user in the center and give them a name and a bit of description about who they are or what they do.

3. Diverge, with each team member writing one observation per sticky note and applying it to the appropriate quadrant of the map.
4. Annotate unknowns (assumptions and questions) for later inquiry or validation.
5. Discuss observations and fill in gaps collaboratively. (Weprin, 2016)

Alternatively, the four quadrants can be labeled Think & Feel, Hear, Say & Do, and See, with additional space for Pain & Gain as recommended by Dave Gray (n.d.), the reputed creator of this idea. Empathy maps can capture the answers to questions such as these:

- What are our respondents' preoccupations, worries, or aspirations?
- What is their attitude in public, appearance, or behavior toward others?
- What are their pain points, their fears, frustrations, or obstacles?
- What are their wants, needs, or measures of success? (Gray, n.d.)

Weprin (2016) suggests that this mapping process be completed as a team and, where possible, that respondents be included in the creation of the map or checking the accuracy of assumptions used to develop the map. He also advises that readers go beyond just listing job titles (e.g., our respondents are teachers or accountants) and consider respondents' "actual tasks, motivations, goals, and obstacles."

Readers of this text may be thinking that this is where we test out some of our question ideas on potential respondents. Some may be familiar with the terms *piloting* or **cognitive interviewing** and think that is what we are describing in this section. Not so. The focus on empathy is squarely on knowing and understanding respondents *in preparation for developing appropriate questions for them*. This is part of the survey design process that *precedes* drafting questions. Pretesting survey questions using various strategies such as those mentioned earlier comes later in the design process (and is discussed in Chapter 7). Finally, empathy can even influence the look, feel, and layout of a survey as well as inform decisions about how and when a survey is administered.

STORIES FROM THE FIELD

WHEN PUTTING RESPONDENTS FIRST YIELDS SUCCESS

Throughout my career in public education, I've worked with school principals for decades and am intimately familiar with their responsibilities and workloads. When I needed the answers to three key open-ended survey questions from a group of 16 busy principals, I knew that an online survey wouldn't do the trick. Even if I thought that composing a few brief paragraphs in response to these questions would only take them 15 to 20 minutes at the most, I knew their world of countless emails, voicemails, and office interruptions on top of their school management and personnel supervision responsibilities well enough to know that sitting down at a computer, clicking on the link, and taking a survey would not rise to the top of the priority list. However, I also knew from previous interactions with them that they cared about the project and that if I asked for that same 15 to 20 minutes of their time for a phone call, they would agree to participate. And they did. My respondent-centered survey design decision resulted in a 100% response rate!

—Sheila B. Robinson

STORIES FROM THE FIELD

EMPATHY IN SURVEY DESIGN

My colleague and I evaluated the services provided by youth centers. As part of the funding requirements, we had to administer a survey to youth who had received services. The survey asked questions about life in general as well as about specific mental health outcomes like substance use, violence, and trauma. Since the survey was mandated by the funder, we could not change the questions. However, we wanted to learn how youth respondents would feel about the survey, so we piloted the survey to youth program staff. Some of the feedback we received was that some of the questions made the staff feel uncomfortable. The youth program staff were worried that some questions could possibly trigger painful memories and may be traumatic to youth responding to the survey. We (the evaluators) attended a training on trauma-informed interviewing. As part of the training, we learned several strategies that we could use to make respondents comfortable such as recognizing and responding appropriately to signs of distress and selecting an appropriate room in which to administer the survey. We also shared these strategies for trauma-informed interviewing with youth program staff. If we had not piloted or talked to the youth program staff, we would not have learned about trauma-informed interviewing. Empathy can be useful not only in designing questions but also in survey administration. An empathetic researcher should be ready to learn new skills or provide resources based on insight gained through empathetic need finding.

—Chithra Adams

The first story in this section illustrates a time when a key survey design decision (in this case, about administration mode) resulted in a positive outcome. This same decision would not necessarily have been made in a different context such as a much larger respondent pool, and the decision was counter to what the researcher initially preferred (an online survey in which respondents typed their own answers). The point is that *the design decision was made based on known characteristics of the respondents* and that was key to the successful outcome. In the second story, empathy was also used to alter the survey administration conditions in a very different context but also in response to a known characteristic of respondents.

Understand

The second phase in the design thinking process is *understand*. In the literature on design thinking and design processes, this phase is also known as *define*.

> The Define mode of the design process is all about bringing clarity and focus to the design space. It is your chance, and responsibility, as a design thinker to define the challenge you are taking on, based on what you have learned about your user and about the context. (Hasso Plattner Institute of Design, n.d.)

After developing empathy for respondents through various efforts to know more about them, we next must make sense of insights gained along the way. In this phase, we reflect on our research or evaluation questions and revisit the clearly articulated survey purpose and information needs in the context of what we now understand about potential respondents. This sense-making phase may also include reviewing relevant literature or even existing surveys on the topic and potential **constructs**, or underlying concepts, we hope to measure. (See Chapter 2 for more on constructs.)

If we have made special efforts to empathize with respondents by observing them in the field or interviewing them, this is where we review, analyze, and interpret notes, transcripts, or other data collected in service to understanding their experiences. If we already have a thorough knowledge of respondent characteristics, this is where we might list, prioritize, and begin to think about what these characteristics will bring to bear on question and overall survey design.

Brainstorm

The third phase in design thinking is *brainstorm*. In some of the design thinking literature, this is also known as ideate, or idea generation. In survey design, this is where we generate and start drafting questions, taking into account what we know about respondents as well as our survey purpose and information needs. As with any good brainstorming session, during this phase "anything goes." What sets this session apart from other types of brainstorming is that we have gone through two previous phases—*empathize* and *understand*—that have grounded us in the experiences of respondents. This positions us well to generate a more focused set of questions than would otherwise be generated without those phases and should help limit necessary pruning of ideas and questions in later stages of the process.

During the brainstorm phase, we are free to generate possible questions with little regard to specific question types. In other words, we don't attend to whether questions will be open- or closed-ended, and we don't necessarily craft rating scales or other types of response options. Whereas designers use the question stem "How might we . . . ?" as they generate ideas for a solution to a perplexing problem, we use it to come up with survey questions. For example,

- How might we ask respondents about their shopping habits?
- How might we find out about their opinions of the new policy?
- How might we understand their experiences at the event?
- How might we figure out if they intend to change their practice?
- How might we ask questions of respondents in ways that limit bias?

Researchers must avoid getting mired in the details of question wording, question order or sequencing, and other design considerations during this phase. New question ideas and considerations may surface while brainstorming that would otherwise be missed if we jump too fast into the details. We strongly encourage survey designers to stay flexible and creative during this phase and avoid rushing the process.

Prototype

The fourth phase of design thinking is *prototype*. A prototype of a product or service is typically designed to be taken to a set of potential users for their feedback and input. In survey design, this phase looks like building a survey draft—our prototype. To do so we draw on what we have synthesized from understanding and brainstorming. Question ideas are refined and then we determine whether they should be open- or closed-ended given the information needed. If closed-ended, specific question types are established, and response options are developed. This is an important early step in ensuring that our survey measures what we really want to measure—that our questions have **face validity**. This requires that the constructs we want to measure are clear and well defined. In other words, we want to **operationalize** them. This typically requires review of relevant existing research about a given concept (e.g., if satisfaction is to be measured, the researcher needs to review and incorporate existing research about how satisfaction is operationalized). At this stage we also begin to ensure that question stems and response options are appropriately aligned (i.e., mirror one another in focus and language; see Chapter 5 for fuller treatment of the development of response options).

During prototyping, questions may be added, or questions generated during the brainstorming phase may be divided into multiple questions. Or questions may be pruned as our understanding of how best to capture the desired information is refined. Questions should also be prioritized based on work from earlier phases to balance information needs with feasibility constraints. In other words, knowing what we know about potential respondents, how long a survey can we create? Generally, fewer questions should be asked than initially brainstormed or drafted. However, it is more important to consider the length of *time* it will take respondents to complete the survey than the *number* of questions included. And because question responses can be influenced by the order in which questions are posed to respondents, it may be useful to build multiple prototypes, or different versions of our survey, in which questions are alternated in order to test this phenomenon. Building multiple prototypes to test may be especially important for larger-scale projects or if the survey includes potentially sensitive questions. In general, we want to introduce more sensitive or potentially threatening questions only *after* more innocuous questions are posed. In addition, we begin surveys with more general questions prior to asking more specific questions. There are also order considerations for **demographic questions**. All of these choices have implications for survey results, and some of these **order effects** are discussed in Chapter 8. Once we receive feedback on a prototype, we may redesign the survey accordingly to incorporate any relevant feedback before moving into a testing phase.

Test

The fifth phase in design thinking is *test*. In survey research, we often refer to this as "pretesting" because we are testing the instrument and its component questions *before* we administer it to our desired respondents. Such testing solicits even more user input to inform potential revision or redesign before a survey is finalized. Common survey research pretesting practices include piloting and cognitive interviewing. For a detailed description of these pretesting strategies, see Chapter 7.

Whether we share survey drafts with a few willing colleagues who offer honest feedback, or invest considerable resources in more formal and larger-scale pretesting, valuable information can be gleaned from this step that cannot be learned any other way. We can ask for reactions and initial impressions, not only of questions but also of format. Although pretesting is clearly time-intensive when thoroughly executed, we cannot overstate its worth in a purposeful survey design process.

An Iterative Process

Design thinking may appear to be a linear process, but it is also an iterative one. Although empathizing is a natural antecedent to understanding, and brainstorming must precede prototyping and testing, a designer frequently returns to earlier phases throughout the process, as more and more information is derived from each phase. As we make sense of our learning during the define phase, we may engage in additional activities to build more empathy for respondents. As we build prototypes or receive feedback from pretesting them, we make revisions based on this new learning. In some cases, this may even cause us to return to the respondent population to seek more of an understanding about them. Being empathetic to respondents means that the researchers are open to possibilities and willing to change, learn new skills, and use new software tools. Design thinking requires flexibility and adaptability. It also means practicing design restraint. Thorough empathy and understanding phases may leave the researcher with a wealth of information about respondents; therefore, it will be important to identify which aspects of respondents' lives will influence how survey questions will be answered and focus on these during the design process.

A PURPOSEFUL SURVEY DESIGN PROCESS

Numerous authors of survey texts have conceptualized a survey research process in a series of stages, steps, or phases. We examined many examples of these before crafting our own process specific to survey question design. In doing so, we noticed the following across the various process frameworks:

- There is no one "right" way to conduct survey research.
- Many authors (including us) at least agree on some *key parts* of the process.
- It is all but impossible to suggest *how long* each part of a survey research process will take, for each survey effort and research project is unique.
- Parts of the design process are interrelated, and the process is typically iterative in practice, despite being presented (as we have) as a more ordered, linear process.

At its most fundamental, we conceptualize the broader survey research process as having four main parts: (1) design, (2) administration, (3) analysis, and (4) use. This text

focuses on the survey design process, that is, the process of developing survey instruments, and particularly survey questions. We consider the design process as inclusive of everything that needs to happen prior to survey administration, including any pretesting of survey questions necessary. Our survey design process is described next and summarized in the boxed outline. This text does not include a great deal of advice about administration, analysis, and use, though considerations are peppered throughout, and Chapter 8 covers preparing for administration, analysis, and use (e.g., determining survey administration mode). Many other survey research texts cover survey administration and data analysis far more thoroughly than we can here. Use of survey results specifically has not been tackled as thoroughly, but many texts support use of research and evaluation results more broadly (e.g., Patton's [2008] *Utilization-Focused Evaluation*).

The survey design process has three major components with a few subcomponents under each.

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

Planning and Predrafting

In this step, we clarify research goals and articulate **research questions**. We determine what will be measured, confirming that a survey is the right tool to capture the data we need. We use design thinking phases of empathizing, understanding, and brainstorming to get to know our respondents in the context of our information needs, and begin a list of more detailed topics for questions that may end up on our survey.

Developing Questions

In this step, we use the valuable information gathered in the planning and predrafting stage to actually select and/or draft survey questions. We determine the extent to which we will need to consider whether a suitable survey instrument already exists or whether we will need to design

an original measurement tool. We use design thinking strategies to prioritize questions brainstormed during the planning and predrafting step and build a prototype of a survey that will be shared in the testing step to determine the comprehensibility and appropriateness of the questions, among other considerations.

Finalizing

In this step, we pretest survey questions using strategies such as peer review, piloting, or cognitive interviewing to ensure that respondents understand questions and can interpret and answer them as we intend. We can also use feedback from respondents to refine and revise questions in preparation for administering the survey. One additional strategy is to use a quality assessment tool, such as our Checklist for Quality Question Design, introduced in Chapter 7.

This step also calls for planning for analysis and use of survey responses and planning for administration details, including mode (e.g., in person or self-administered, telephone, mail, or online) and timeline. These decisions are based on what we know about respondents and the context in which the survey is being administered. We should determine how results can (and cannot) be used and consider what analyses will be needed. Ideally, we begin considering administration, analysis, and use intentions and needs far before we are done developing a survey tool. This is a great example of how the survey design process can be highly iterative rather than linear; analysis plans may evolve over time as the survey is developed and the researcher determines what types of questions are needed.

AN ITERATIVE PROCESS

Though the process appears linear (and our text roughly follows the order presented), it is critical to note that the entire process is typically iterative. For example, though a researcher may begin with determining survey purpose, then develop questions, and then employ a pretesting strategy, pretesting may reveal that she needs to know more about respondents and their specific characteristics or context and rework some questions before returning to conduct additional pretesting. Additionally, there are a few pieces in particular that are hard to place in a linear order. Understanding respondents is often an evolving effort, and researchers are wise to do at least some planning (or at least *thinking*) about administration and use of results early in the process.

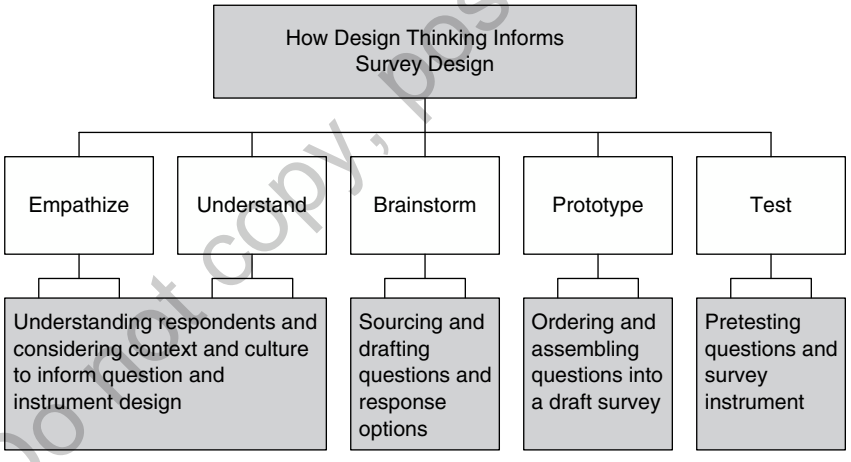
Table 1.1 shows the stages of our survey design process and where readers can find more information about these stages in this book.

We intentionally introduced design thinking before introducing our survey design process. Understanding design thinking will impel the reader to keep the survey design process focused on respondents, which is critical to crafting effective surveys. Design thinking can be thought of as a set of fluid phases that, in a sense, *overlay* the survey design process and inform our thinking throughout (see Figure 1.2).

TABLE 1.1 ■ Purposeful Survey Design Processes

Purposeful Survey Design Process	Location in Text
1. Planning and Predrafting	
a. Determining and articulating survey purpose	Chapter 2
b. Understanding what surveys can measure	Chapter 2
c. Understanding survey respondents	Chapter 3
2. Developing Questions	
a. Sourcing questions	Chapter 4
b. Crafting question stems and response options	Chapters 4 & 5
c. Considering demographic and other sensitive questions	Chapter 6
3. Finalizing	
a. Pretesting	Chapter 7
b. Preparing for administration, analysis, and use	Chapter 8

FIGURE 1.2 ■ How Design Thinking Informs Survey Design



DISCUSSION QUESTIONS

- What are the implications of survey fatigue for data collection using surveys? How might researchers work to limit survey fatigue?
- What phase of the design thinking process seems easiest to apply to survey design? Most challenging? Where might a survey researcher devote the most time and effort?

- How can understanding your desired respondents help in designing quality survey questions? How might a survey researcher access information needed to understand desired respondents?
- What are the advantages of the prototyping and testing phases?

DESIGN DRILLS

1. Find a question in a publicly available survey (such as a retail customer service survey or a student survey from your school) that is problematic. How is it problematic? What do you think went wrong in its design? Are any of the reasons for purposeful survey design outlined in this chapter illustrated by this question? If so, how?
2. Review the following questions and identify the problem. Rewrite the question with necessary improvements (making any assumptions necessary about what information might be desired by the researchers).
   a. Was the information you received timely, relevant, and appropriate?
   _____YES
   _____NO
   b. Do you think it's wise to cut spending on Program A and increase spending on Program B?
   _____YES
   _____NO
   c. How useful was this program for you?
   _____Excellent
   _____Very good
   _____Fair
   _____Poor
   d. What is your age?
   _____under 30
   _____30–35
   _____35–40
   _____40–45
   _____45–50
   _____over 50
3. Research Scenario, Part 1 (additional parts are found in subsequent 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 geographic region. Assuming for a moment that a survey of participants is the best way to capture critical data for this project, draft a brief memo or email to your research teammates or supervisor making a case for a purposeful survey design process. Include a proposal for how you might work

to understand your desired respondents as part of that process, noting why that will be important for this project.

4. Your survey: In this chapter we learned that designers use the question stem "How might we . . . ?" as they generate ideas for a solution to a perplexing problem. We provided examples of how this might be used to help draft survey questions. For example, "How might we ask respondents about their shopping habits?" Or "How might we ask questions of respondents in ways that limit bias?" For a survey you are currently developing, or expect to develop, generate a few additional "how might we" questions specific to your research scenario that could help you draft questions with a respondent-centered mindset.

EXTENDED LEARNING

- A thorough treatment of sampling can be found in the following:
  - Alreck, P. L., & Settle, R. B. (2004). *The survey research handbook* (3rd ed.). Irwin.
  - Babbie, E. (1990). *Survey research methods* (2nd ed.). Wadsworth.
  - Ornstein, M. (2013). *A companion to survey research.* Sage. (Provides a good overview of applied sample design and the design of purposive samples.)
  - Rea, L. M., & Parker, R. A. (2014). *Designing and conducting survey research: A comprehensive guide* (4th ed.) Jossey-Bass.
  - Weisberg, H. F., Krosnick, J. A., & Bowen, B. D. (1996). *An introduction to survey research, polling, and data analysis* (3rd ed.). Sage.
- Additional details about survey error can be found in Czaja, R. F., & Blair, J. (2004). *Designing surveys: A guide to decisions and procedures* (2nd ed.). Sage. Dillman, D., Smyth, J. D., & Christian, L. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method* (4th ed.). Wiley.
- For more on use-focused evaluation, see Patton, M. Q. (2008). Utilization-focused evaluation. Sage. Patton, M. Q. (2011). *Essentials of utilization-focused evaluation*. Sage.
- For more on design thinking:
  - Tim Brown's blog: *Design thinking*, http://designthinking.ideo.com
  - *The field guide to human-centered design*, http://www.designkit.org/resources/html
  - Cameron Norman's blog: *Censemaking: Contemplating complexity, designing social innovation*, https://censemaking.com
  - Hasso Plattner Institute of Design, *An introduction to design thinking process guide*. https://web.stanford.edu/~mshanks/MichaelShanks/files/509554.pdf