CHAPTER 13

Survey Research

James F. Baumann
James J. Bason

What are elementary students' attitudes toward reading in school? What do middle school students choose to read in their free time outside school? How do high school teachers accommodate students who struggle to read and understand content textbooks? How do parents support at home the reading instruction their children receive at school? What content do university teacher educators cover in their elementary reading methods courses? What are the teaching experiences and knowledge base about reading of upper-level school district administrators? Answers to questions such as these can be addressed through survey research.

It is the purpose of this chapter to describe how surveys can be used by literacy researchers to address a variety of questions about the characteristics of educational groups. We begin with a definition of survey research, followed by a brief history of literacy-related survey research. Next, we describe the various types of surveys, and then we address how survey researchers identify populations, draw samples, and make inferences from survey data. Next, we describe the process in which researchers engage when implementing a survey inquiry, followed by a discussion of quality standards for survey research. We conclude by presenting two exemplars of survey research projects and a summary of the chapter.

WHAT IS SURVEY RESEARCH?

According to the American Statistical Association, *survey* "is used most often to describe a method of gathering information from a

sample of individuals" (*What Is a Survey?*, n.d., ¶ 3). Jaeger (1997) indicates that "the purpose of survey research is to describe specific characteristics of a large group of persons, objects, or institutions" (p. 449), which led him to the colloquial definition of *survey research* as a "study in which data are collected from part of a group, for the purpose of describing one or more characteristics of the whole group" (p. 450). Surveys are popular methods of collecting information from individuals and the preferred means to address a research question when it is most efficient to simply *ask* those who can inform the question. A survey typically involves the administration of a questionnaire or interview to a relevant group.

One way to describe surveys involves the degree of organization they entail. In this chapter, we focus on *structured surveys*, that is, research tools that involve having persons respond to a series of questions (and perhaps follow-up questions) through various media. Structured surveys generate numerical data directly or data that can be categorized and tabulated such that the data can be explored quantitatively through descriptive or inferential statistics.

In contrast, researchers can engage in *semistructured interviews* or *qualitative interviews*, in which the interviewer and interviewee engage in more of a conversation than in an interview. Qualitative interviews generate very rich and in-depth data as opposed to structured interviews, and hence they address different research questions. We confine our discussion to structured interviews in this chapter and refer readers to other sources for information about semistructured or qualitative interviews (Gubrium & Holstein, 2002; Seidman, 1998).

Finally, it should be understood that surveys are not limited to humans. For example, a researcher might survey objects such as instructional materials or achievement records. Most educational survey research, however, has involved questions about the attitudes, knowledge, experiences, and behaviors exhibited by *persons*: teachers, students, administrators, parents, policymakers, or others interested in education. Thus, we focus on survey research that has the potential to address research questions that describe human involvement in literacy education.

A BRIEF HISTORY OF SURVEY RESEARCH IN LITERACY

The beginnings of survey research can be traced back to the late 19th century, with the modern era of sample surveys beginning in earnest in the early 20th century (Weisberg, Krosnick, & Bowen, 1996). The

first survey research in literacy occurred in the early 1900s and was tied to what Venezky (1984) referred to as "school surveys," in which U.S. school districts evaluated their entire educational programs, including reading instruction. For example, Venezky described a survey in Cleveland as involving "analyses of the goals of reading instruction, the training of teachers, the methods and materials and time used, and the achievements of students in different components of reading" (p. 18).

Survey research in the social, behavioral, and educational sciences gained widespread acceptance in the 1950s, becoming one of the most popular and efficient methods of collecting information from individuals (Weisberg et al., 1996). In 1961 Austin, Bush, and Huebner described in detail the process of "The Reading Survey" as a "survey method for getting a complete perspective on the reading program" (p. 131), and several significant reading survey studies were conducted during this period. Austin and Morrison (1961) published *The Torchlighters*, a mail survey of reading teacher education programs at 74 U.S. colleges and universities, which was replicated 13 years later (Morrison & Austin, 1977). As a follow-up to *The Torchlighters*, Austin and Morrison asked in 1963, "what guidance do teachers receive *after* they complete their baccalaureate education?" (p. ix). They addressed this question in *The First R* (Austin & Morrison, 1963) by conducting a mail survey of administrators in 1,023 U.S. school districts about the content and conduct of reading instruction and by conducting follow-up, on-site visits in 51 school districts. Barton and Wilder (1964) also reported an extensive study of reading that included face-to-face interviews with leading experts in reading education and mail surveys to university faculty, teachers and principals, and the general public.

Survey research has flourished from the 1970s to the present (Weisberg et al., 1996), with a number of recent surveys in the field of literacy education. For example, researchers have explored elementary students' attitudes toward reading (McKenna, Kear, & Ellsworth, 1995), middle school students' motivations for reading (Ivey & Broaddus, 2001), professional development for literacy educators (Commeyras & DeGroff, 1998; Hughes, Cash, Klingner, & Ahwee, 2001), reading experts' opinions of contemporary trends and practices (Baumann, Ro, Duffy-Hester, & Hoffman, 2000; Flippo, 1998), and educators' views of state-mandated, high-stakes testing (Hoffman, Assaf, & Paris, 2001), along with modified replications of Austin and Morrison's *The Torchlighters* (Hoffman & Roller, 2001) and *The First R* (Baumann, Hoffman, Duffy-Hester, & Ro, 2000).

TYPES OF SURVEYS

Surveys come in various forms and can be organized in different ways. The most common way to characterize surveys is by the method of data collection. Survey collection methods include face-to-face surveys, telephone surveys, mail surveys, and technology-enhanced survey methods. The decision regarding which method to use is typically based on the nature and size of the population under study, the content of the information to be collected, the length of the survey, or the difficulty of the task respondents are asked to complete.

Face-to-Face Surveys

Face-to-face surveys, or structured personal interviews, are preferred by many survey researchers because they afford the most flexibility, allowing a researcher to control the dynamics of the interview process considerably. Jaeger (1997) noted that "apart from their costs, face-to-face interview surveys have so many advantages that some survey researchers consider alternative methods to be totally unacceptable" (p. 459). Advantages of face-to-face surveys include the opportunity for the interviewer to ask for clarification, to ask follow-up or "branched" questions, and to observe and make note of surroundings. In addition, face-to-face surveys tend to result in high response rates. When depth of information and a flexible format are desired and feasible, a face-to-face survey may be the preferred method. When a researcher is interested in a rich, interactive description by a group of individuals on a specified topic, a *focus group* may be employed. However, given the unique purposes and requirements of focus groups, we do not pursue them further in this chapter and refer interested readers to other sources (Bloor, Frankland, Thomas, & Robson, 2001; Greenbaum, 1998; Krueger & Casey, 2000; *What Are Focus Groups?*, 1997).

Telephone Surveys

When many interviews need to be conducted or the sample of prospective respondents is geographically dispersed, the preferred method may be a *telephone survey*, which is administering a questionnaire during a phone conversation. Telephone surveys retain some of the advantages of face-to-face surveys—for example, the administrator can explain the purpose of a survey and ask follow-up or branched questions—while providing for a more economical way to gather information. In addition, a telephone survey researcher can

inquire why an individual chooses to participate in a survey or not. Organizations such as the Harris Poll and the Gallup Poll conduct telephone interviews, as do marketing firms and researchers in the social, behavioral, and educational sciences.

A common approach to identifying telephone numbers and optimizing response rates is to employ the *list-assisted method*. In this approach, research firms combine published telephone directory information and random selection of numbers, resulting in telephone lists that have response rates between 50% and 85% (Kalton, 1983; Nachmias & Nachmias, 1981). The list-assisted method may be combined with *random digit dialing*, in which computers randomly generate telephone numbers such that every household in a given area (even those with unlisted numbers) has a known probability of being selected, resulting in highly accurate samples being drawn from specific populations (Tucker, Lepkwoksi, & Piekarski, 2002).

Telephone surveys are no panacea, however. Increasing use of telephone surveys among the general population and increased telemarketing and fundraising activities using the telephone have contributed to increasing refusal rates and growing concern about *nonresponse bias*, that is, failure to represent the views of those who choose not to participate (Groves, Dillman, Eltinge, & Little, 2002). The proliferation of cell phones, which are typically excluded from survey listings, and the use of telephone answering and caller ID devices used to screen incoming telephone calls raise additional concerns about bias using the telephone survey method (Link & Oldendick, 1999). In spite of these limitations, it is generally agreed that the telephone remains one of the most effective ways to collect data from large samples of individuals.

Mail Surveys

When it is not feasible to conduct face-to-face or telephone surveys, a *mail survey*, distributing a written questionnaire through the postal service, is a viable alternative. Advantages of mail surveys include relatively low cost, the ability to access a broad sample of potential respondents, and the possibility of obtaining a large number of responses. In addition, there is evidence that mail surveys tend to minimize *social desirability bias* (respondents answering in ways deemed to be socially acceptable or appropriate rather than reflecting their actual attitudes or behaviors) when compared to other survey methods (Hochstim, 1967; Wiseman, 1972).

Mail surveys are not without disadvantages, however. While economical, they may not produce as a high a response rate as other

types of surveys, one must rely on the respondent to interpret questions properly (i.e., no opportunity for clarification), a researcher has no control over the actual administration of the survey, and one cannot determine why given sample members choose to participate or not participate. Most mail surveys employ the total design method (Dillman, 1978) to reduce nonresponse bias. The total design method uses an initial mail survey followed by one or more waves of follow-ups to nonresponders to enhance participation and minimize nonresponse bias. Typically, after the first survey is mailed, a postcard follow-up reminder is sent to all sample members, followed by the mailing of another questionnaire to nonresponders. Response rates to mail surveys vary widely depending on the amount and effort at follow-up.

In spite of their limitations, "mail surveys are a powerful, effective, and efficient alternative to their more expensive relatives—the telephone survey and the personal interview" (*More about mail surveys*, 1997, p. 1). Indeed, mail surveys have been (e.g., Austin & Morrison, 1961, 1963) and remain (e.g., Baumann, Hoffman, et al., 2000; Baumann, Ro, et al., 2000; Commeyras & DeGroff, 1998; Hoffman et al., 2001; Hoffman & Roller, 2001; Hughes et al., 2001) one of the most popular methods employed in literacy survey research.

Technology-Enhanced Survey Methods

A new set of data collection methods relying on technology include web- and Internet-based surveys. *Web surveys* require a respondent to visit a website, enter a password to access a survey instrument, and complete a survey online (Couper, 2002). *Internet surveys* involve sending an e-mail message to prospective respondents and inviting them to participate in a survey by completing a questionnaire in the body of the e-mail (Schaefer & Dillman, 1998).

Despite the advantages of low cost and speed provided by web- and Internet-based surveys, such surveys are limited to special populations known to have high rates of Internet access and thus tend not to represent broader populations. There currently is no known reliable means to access e-mail addresses, so it is difficult to obtain representative samples. These surveys are effective for known Internet users for whom e-mail addresses are readily accessible (e.g., literacy professors whose e-mail addresses are contained in university directories), but this requires the creation of specialized sample lists. With increasing Internet access in households and the growing availability of commercially available e-mail lists, however, it is likely that web-

and Internet-based survey use and credibility will increase in the future.

Technology is also enhancing the application of more conventional survey procedures. *Computer-assisted telephone interviews* are becoming increasingly common, in which the interviewer reads questions from a computer screen and uses a keyboard to enter an interviewee's responses directly, enhancing the speed and reliability of telephone interview data. Similarly, one finds face-to-face surveys being administered by interviewees who use notebook computers to conduct *computer-assisted personal interviews* (Couper et al., 1998)

With the advent of computer programs that can recognize speech, telephone surveys can be administered through *interactive voice recognition*, in which a respondent calls a toll-free number and communicates directly by saying prescribed responses. Similarly, using *touch-tone data entry*, interviewees proceed through a prerecorded interview in which they are asked to enter responses through their telephone touch-tone key pad. These approaches are most useful when factual information is required from the respondent, or when respondent–interviewer interaction is not a necessary requirement for completing an interview.

POPULATIONS, SAMPLES, AND INFERENCES

When a researcher's question requires responses from a small number of persons, for example, learning about the content area reading practices of subject-matter teachers at a particular high school, then it makes sense to interview all individuals. This would involve a census of the entire *population*, or surveying all persons who fit a particular classification. A *population parameter* is a value that describes an entire population, for example, the *mean*, or average, number of years of teaching experienced for the population of all full-time, public grade 10 English teachers in a specific school district.

When the population under study is large, however, for instance, when Phi Delta Kappa annually polls Americans about their attitudes toward public schools (Rose & Gallup, 2002), a census of the entire population is impractical, costly, or simply not possible. In this situation, other procedures are used to select a subset of the entire population. In fact, most surveys involve a subset of individuals, or *sample*, and most survey researchers wish to gather data on a sample in order to generalize to a broader group, that is, to estimate a population parameter.

To be able to make generalizations from a sample to a population, a survey researcher must begin by specifying a *sampling frame*, which includes all the members of a given population. Frames or populations might be small or large. For example, large populations would include all full-time, public school kindergarten teachers in the United States or all school district superintendents in the 50 states. Examples of smaller populations would be all parents of K–5 students in a small, rural school district or all students who have completed an academic assistance course in reading within the past year at a small state university.

Once a sampling frame is specified, a researcher must determine how to select a *probability sample*, or a subset of the specified population to interview and from which one can make generalizations. Two common probability samples are a simple random sample and a stratified random sample. A *simple random sample* is one in which everyone in the sampling frame has an equal chance of being selected through some random process, such as through the use of computer random sampling or a random table of numbers. For example, randomly selecting 1,000 names from all full-time K–5 classroom teachers in the state of Georgia would be a simple random sample. A *stratified random sample* involves selecting participants from certain subgroups, or strata, within a sample. For example, if there were smaller class sizes in the primary grades, there would be proportionally more teachers at those levels. In this situation, a researcher might identify each grade level as a stratum, or subgroup, and then randomly sample each grade level proportionate to its representation in the overall population. This would result in a sample that approximates the overall K–5 population better than a simple random sample.

When engaging in probability sampling research, any *sample statistic*, or value that describes the sample such as a sample mean, is not exact; there is some error associated with it. For instance, the average class size of a simple random sample of 1,000 full-time second-grade classroom teachers in South Carolina would be somewhat different from the actual population parameter should one be able to poll all teachers. If a researcher wishes to use sample statistics to estimate population parameters, the researcher needs to know and report the *sampling error*, or an estimate of the degree of error associated with a sample statistic. This estimate is commonly referred to simply as the "margin of error" (*What Is a Margin of Error?*, 1998).

The margin of error in a sample statistic is usually reported in plus or minus terms and at some degree of confidence. For example, Baumann, Hoffman, et al. (2000) reported that 1,207 responses to their teacher survey resulted in a 2.7% sampling error at the 95%

confidence level. This means that any sample statistic they reported for the full sample of teachers would represent the "true" or population parameter 95% of the time within plus or minus 2.7% of the reported score.

The sample margin of error is directly related to sample size. The larger the sample, the smaller the margin of error, the more the sample statistic approximates the population parameter, and the more confidence a consumer can have in the precision of survey results. For example, a simple random sample of 100 would result in about a 10% margin of error, a sample of 500 about 4.5%, and a sample of 1,000 about 3% (*What Is a Margin of Error?*, 1998, p. 6). A larger sample size is always preferable, but from a cost and efficiency perspective, there often comes a point at which collecting additional cases is not worth the per-case cost. Once sample sizes reach about 1,200, the reduction in sampling error becomes smaller as more and more cases are added, which is why one often finds sample sizes of around 1,000 for public opinion polls.

One final point: The margin of error as we have discussed it refers only to the sampling error. There are other possible errors in survey research, however. These *nonsampling errors* include poorly worded or confusing questions, nonresponse bias, and social desirability bias, to mention a few. Such nonsampling errors cannot be estimated mathematically, so a survey research consumer should be aware of these other sources of error (Fowler, 2002) and read an article critically to determine the degree to which the researcher has attempted to control for, or at least acknowledge, these additional sources of error.

SURVEY RESEARCH DESIGN PROCESS

Engaging in any scientific inquiry involves a careful, methodical process. Although there are no uniform, agreed-on number of steps in planning and conducting a survey research project, we present a six-step process that captures the essence of recommendations offered by other survey research methodologists (e.g., Jaeger, 1997).

1. *Formulate a research question and hypotheses.* If one accepts the notion that quality research begins not with methods but a carefully framed question (Shulman, 1997), then it follows that the selection of survey research as a method must depend on the formulation of a core research question that is amenable to being answered through a survey. Questions such as those we used to open this chapter are of

the type that would be appropriate for a survey study. Theory should be used to guide the development of the research question, and a thorough literature review of past and current research findings should be conducted to craft specific research hypotheses.

2. *Devise a measurement strategy.* In this step, a researcher must determine the type of survey and specific items that will best measure the features of the underlying psychological, behavioral, social, or educational topic or issue under consideration, or what is commonly referred to as *construct validity.* Special care also should be taken to ensure that measurement error is minimized. Although it is impossible to eliminate all measurement error, several sources of error can be reduced, including sampling, nonresponse, interviewer, and respondent error. Sampling error is probably the easiest source of error to control. If probability sampling procedures are used, the range of potential error is known by way of the sampling margin of error. Nonresponse error is more difficult to control. The lower the effective response rate in a survey, the greater the likelihood that nonresponders may be systematically different from responders. A researcher needs to obtain additional information about nonrespondents in order to determine if responders and nonresponders differed in the way they answered survey items. Interviewer error occurs when interviewers fail to follow standardized procedures, for example, not reading survey questions exactly as written. Training survey interviewers thoroughly minimizes interviewer errors. Finally, respondent errors occur when survey items are unclear; when respondents engage in *satisficing,* or selecting an answer considered to be reasonable to the researcher (see Krosnick, 1999); or when respondents provide answers to survey questions even when they may not actually have an opinion on the topic (Converse, 1964). A researcher has responsibility for constructing clearly worded items and for being aware of, minimizing, or at least acknowledging, other respondent error.

3. *Determine the sampling frame and data-collection method.* The next step is to select the appropriate sampling frame, which will be determined, in large part, by the nature of the research question. If one wishes to seek information from teachers, for example, a researcher must determine which teachers to sample: their position type, the grade levels they teach, their subject specialization, the type of schools in which they work (e.g., public vs. private), and the like. A researcher can make generalizations only from the limits of the sampling frame, so identifying the appropriate group to survey is essential. Once the relevant population has been specified, the researcher must determine which method will be used to collect survey data—

that is, whether to choose a face-to-face, telephone, mail, or technology-based survey. Again, the research question, as well as pragmatics such as research budget and the availability of sampling sources (e.g., professional organization lists and marketing research lists), will generally guide a researcher in selecting the most suitable survey type.

4. *Establish sampling methodology and sample size.* The next step in the design process is to determine the method for sampling. If one is surveying a small population, it may be feasible to try to secure responses from all individuals in the population. If one is sampling a large population, which is typical in survey research, one must determine whether the sample will be a selective or a probability sample. If the sample is selective—for example, selecting parents who are members of a local parent–teacher organization—one cannot make generalizations to a broader population of parents. There may be times at which a selective sample is appropriate or when it is the only feasible option, but then a researcher must recognize that generalizing to a larger group is not possible. When a researcher employs a probability (i.e., random) sample, which is the typical standard in survey research, then the researcher should make a decision about sample size with the understanding of what the margin of error will be. When economically and pragmatically feasible, a sample should be sufficiently large in order to keep the margin of error reasonably low, so that statements about how a sample may represent the population under study can be made within appropriate confidence limits. When smaller random sample sizes must be selected, a researcher must understand the limits this imposes on generalization and acknowledge those in a research report.

5. *Collect and analyze data.* Data collection should be systematic, thorough, and consistent with the accepted procedures for the type of survey employed. For example, if a mail survey were used, the researcher should follow a series of phases in distributing, tracking, and monitoring surveys to promote an acceptable response rate and minimize nonresponse bias (Dillman, 1978). Likewise, data analysis procedures should be thoughtful, consistent with the research questions, and systematic. One must establish a data management and analyses plan prior to data collection, ideally when a study is conceived, so that analyses adequately test research hypotheses. In some cases, simple descriptive statistics such as cross-tabulations will be sufficient and appropriate to answer research questions. In other cases, particularly when a probability sample is obtained, inferential statistics may be appropriate in order to make generalizations from the data. Procedures such as factor or path analyses are commonly used in analyzing survey data. When inferential statistics are employed,

attention should be paid in advance to sample size to ensure that adequate statistical power will be present for the analysis method to be used (Kraemer, & Thiemann, 1987). In short, careful planning at the front end of the survey process will ensure that analysis requirements are met satisfactorily.

6. *Report findings*. Survey researchers ought to have an intended audience in mind as a study is conceived and planned, and this same audience should be considered and addressed as a report is written. The information contained in the following section and accompanying table can also guide a survey researcher in preparing a research report. Although the form, style, voice, and level of detail may vary depending on one's audience and intended publication outlet (e.g., a research journal vs. a more applied serial), basic standards for reporting an inquiry should be followed. The fifth edition of the *Publication Manual of the American Psychological Association* (2001) is a commonly used standard for writing for publication. And as usual, Strunk and White's (2000) classic "little book," *The Elements of Style* (4th ed.), is always a useful tool for a writer who wishes to be concise and precise.

QUALITY STANDARDS FOR SURVEY RESEARCH

Consumers of survey research should be able to examine key components of a data-collection effort to ensure that quality standards have been used, and it is the responsibility of survey researchers to ensure that they provide essential information to allow consumers to effectively evaluate a given survey effort. As a general rule, the methods and procedures used in a survey effort should allow for replication of the study by other researchers. To accomplish this, the population under study should be clearly defined, and the methods and procedures used to conduct the study should be clearly outlined. This includes how the sample was drawn, the exact question wording of the survey (which is often reproduced in an appendix to a report), the number of cases collected during the study, response rate information, and the sampling margin of error. Useful information about evaluating response rates, cooperation rates, and refusal rates for a study can be found through the American Association of Public Opinion Research (*http://www.aapor.org*) and the Council of Academic Survey Research Organizations (*http://www.casro.org*). Both of these professional associations provide valuable information on best practices for survey research, formulas to calculate response and refusal rates, and other helpful information to evaluate survey research quality.

As a further aide to evaluate survey research, we reproduce Jaeger's (1997) survey evaluation checklist in Table 13.1. Although Jaeger has prepared this as a means to help research consumers evaluate published studies, the checklist may also be useful to survey researchers as they conceive of and implement studies of their own.

LITERACY SURVEY RESEARCH EXEMPLARS

To illustrate the survey-research process in literacy, we have selected two exemplars: the McKenna et al. (1995) national survey of children's attitudes toward reading and the Baumann, Hoffman, et al. (2000) national survey of elementary reading instructional practices. Table 13.2 contains a consolidated version of Jaeger's (1997) survey evaluation criteria and presents a summary of how each of these literacy survey studies compares to these standards.

A Survey of Children's Attitudes toward Reading

McKenna et al. (1995) argued for the importance of understanding reading attitude in order to promote reading achievement and to combat aliteracy, and they asserted that research on reading attitudes had been plagued by ill-defined theories, the use of instruments with poor psychometric qualities, and inadequate samples (see Henk & McKenna, Chapter 10, this volume, for further discussion). They provided a detailed review of the theoretical literature on reading attitude and described the development of models by Matthewson (1994) and McKenna (1994), the latter of which was used to guide their investigation and which led them to pose three research questions:

1. What are the overall developmental trends in recreational and academic reading attitude across the elementary grades?
2. What is the developmental relationship between recreational and academic reading attitude, on the one hand, and (a) reading ability, (b) gender, and (c) ethnicity, on the other?
3. What effects on reading attitude can be ascribed to the use of basal reading materials? (McKenna et al., p. 942)

Participants were children in grades 1–6 U.S. classrooms. Their sampling frame was constructed by contacting officers of state and local councils of the professional organization, the International Reading Association, who identified potentially participating school

TABLE 13.1. A Short Checklist for Survey Evaluation

1. Does the report contain a list of specific research questions or issues the survey is intended to address?
2. Do the research questions posed by the investigators appropriately and adequately address the topic of the survey; e.g., in a survey on poverty in the United States, does the research include an examination of poverty as a function of race, level of education, and geographic location?
3. Are the research questions posed by the investigators well organized and well structured?
4. Does the report identify the target population to which generalization was desired?
5. Does the report describe available sampling frames?
6. Does the report indicate a close match between the target population and the operational population?
7. Does the report describe the sampling procedures used? Were probability sampling procedures used?
8. Are nonresponse rates reported for the entire survey and for individual questions?
9. Were nonresponse rates low enough to avoid substantial bias errors?
10. Are any analyses of potential sampling bias reported?
11. Are sample sizes sufficient to avoid substantial random errors? Are standard errors of estimate reported?
12. Is the primary mode of data collection (i.e., mailed questionnaires, telephone interviews, face-to-face interviews) consistent with the objectives, complexity, and operational population of the survey?
13. Are survey instruments provided in the report?
14. Are instructions for completing the survey clear and unambiguous?
15. Are questions on instruments clear and unambiguous?
16. Do questions on instruments encourage respondents' honesty in admitting lack of knowledge or uncertainty?
17. Are questions on instruments free from obvious bias, slanting, or "loading"?
18. Was the survey consistent with ethical research practice; e.g., was the anonymity and/or confidentiality of respondents protected?
19. Does the report contain a description of field procedures?
20. Are field procedures adequate and appropriate? Is it likely that major sources of bias error have been avoided?
21. Are data analyses clearly described?
22. Are data analyses appropriate to the purposes of the survey?
23. Did the survey provide answers to the research questions posed by the investigators?
24. Are the researchers' conclusions sound, or are alternative interpretations of findings equally plausible?
25. Does the survey report contain descriptions of deviations from plans for survey implementation and the likely consequences of such deviations?
26. Does the survey report contain an analysis of the quality of the survey?

Note. From Jaeger (1997, pp. 475–476). Copyright 1997 by the American Educational Research Association. Reproduced by permission.

TABLE 13.2. Evaluation Summary for the Survey Research Exemplars

Jaeger's (1997) criteria[a]	Studies critiqued	
	McKenna et al. (1995)	Baumann, Hoffman, et al. (2000)
Research questions (1–3)[b]	• Theoretically based, explicit research questions posed.	• Theoretically and empirically based research question.
Sampling and response rates (4–11)	• Large (18,000+) national, stratified sample. • Response rates and sampling biases analyses not discussed. • Sample margin of errors not reported, but survey reliabilities acceptable.	• Probability sample for primary (teacher) population, with linked samples for administrators. • Overall survey response rates provided but not item-by-item rates. • Sample margin of errors reported for each survey.
Instrumentation (12–18)	• Survey appended to article. • Reliability and validity data provided.	• Teacher survey appended to article; administrator surveys available from authors. • No reliability and validity data provided.
Data-gathering procedures (19–20)	• Data-gathering procedures summarized.	• Field procedures described in detail.
Analyses, conclusions, report (21–26)	• Descriptive statistics provided in tabular form. • Inferential analyses addressed research questions. • Detailed discussion with plausible interpretations of findings.	• Descriptive statistics integrated into article narrative. • Thorough discussion of findings and research limitations.

[a]Adapted from Jaeger's (1997) full survey evaluation checklist presented in Table 13.1.
[b]Numbers correspond to Jaeger's 26 questions presented in Table 13.1.

districts. The geographically diverse sample, stratified on the basis of gender and ethnicity, included 18,185 children from 229 schools in 95 school districts in 38 states. A table presented the breakdown of the sample according to reading ability, grade, gender, and ethnicity, and a figure presented a U.S. map that showed sampling by state.

The reading attitude instrument used was the McKenna and Kear (1990) Elementary Reading Attitude Survey, a "20-item, 4-

node, pictorial rating scale . . . based on the cartoon character Garfield . . . [that comprised] two 10-item subscales for recreational and academic (school-related) reading attitude" (McKenna et al., 1995, p. 943). Reliability data (coefficient alphas ranging from .74 to .89) and validity evidence (through factor analyses) were provided.

Surveys were mailed to participating teachers along with detailed administration instructions, which included making it clear to students that there were no correct answers. To preclude limited decoding ability from confounding attitude measurement, the teachers read each item aloud twice as students followed along. Teachers were asked to indicate the degree to which they relied on basal readers and to note each student's gender and ethnicity, along with their best professional judgment as to whether each student was above average, average, or below average in reading ability.

Descriptive statistics were presented through tables of means and standard deviations for the recreational and academic reading scales. Inferential statistics involved a series of one-way and factorial analyses of variance that were complemented by line graphs depicting data trends. The authors summarized their findings in five points that correspond to the research questions and subparts as follows:

1. Recreational and academic reading attitudes, on average, begin at a relatively positive point in Grade 1 and end in relative indifference by Grade 6.
2. Increasingly negative recreational attitude is clearly related to ability, and the trend is most rapid for least able readers. The attitudinal gap among ability levels widens with age. For academic reading attitude, however, the negative trend is similar regardless of ability.
3. Girls as a group possess more positive attitudes than boys at all grade levels, both toward recreational and academic reading. In the case of recreational attitude, this gap widens with age. In the case of academic attitude, it remains relatively constant. These gender differences appear to be unrelated to ability.
4. Ethnicity appears to play little role in the negative trend in either recreational or academic reading attitude.
5. The extent of a teacher's reliance on basal readers does not appear to be meaningfully related to recreational or academic reading attitude. (McKenna et al., 1995, pp. 951–952)

The authors interpreted these findings as general support for the McKenna (1994) model of reading attitude, noting that the strongest support came from data indicating that the relationship between reading attitude and reading ability strengthens across the elementary school years. According to McKenna et al. (1995), the results

support "a model of reading in which social factors and expectations gradually shape attitudes over time" (p. 935). The authors call for further research on reading attitudes that employs observational rather than self-report methods and for studies that explore the impact instructional techniques have on students' reading attitudes.

A Survey of Reading Instruction Practices

Baumann, Hoffman, et al. (2000) conducted a modified replication of Austin and Morrison's 1963 *First R* study of U.S. public school elementary reading instruction. Noting that debate over the quality and status of elementary reading instruction preoccupied literacy professionals, policymakers, and the general populace in the late 1990s—just as it had in the early 1960s (Morrison, 1963)—the authors argued that a contemporary, empirical benchmark was needed to address the research question: "What is the nature of elementary reading instruction practices today" (Baumann, Hoffman, et al., 2000, p. 342). Following a search for and reconstruction of original survey instruments, the researchers constructed three forms of a mail survey (classroom teacher, building administrator, and district administrator form) that included multiple-choice, Likert, and open-response items that permitted cross-decade comparisons and allowed for querying educators about contemporary issues.

The primary sample was a probability sample of 3,199 elementary teachers drawn from a national listing of over 900,000 teachers. Modeling *First R* procedures, the building and district administrator samples were linked to the teacher sample, with teacher respondents identifying 623 building administrators, who, in turn, identified 91 district administrators for sampling. Employing the total design method (Dillman, 1978), surveys were distributed, collected, and analyzed by a university survey research facility. Response rates and sampling errors (at the 95% confidence level) were reported for the teacher (37.7% and 2.7%, respectively), building administrator (25.8% and 7.6%), and district administrator (52.7% and 14.1%) surveys.

Quantitative analyses involved the examination of descriptive statistics and selected cross-tabulations. A subset of teacher open-response items and all administrator open-response items were analyzed systematically for themes and trends. Results, reported as percentages of respondents, were juxtaposed to *First R* data by categories, permitting a historic and contemporary look at elementary reading education practices.

Results revealed similarities between teachers of the 1960s and 1990s in that both tended to teach self-contained, heterogeneous

classes of students; provided significant time for reading instruction; taught phonics explicitly; were generally not satisfied with their university training in reading instruction; administered required standardized tests; and were challenged by teaching underachieving readers. Differences were noted in the areas of instructional philosophy (a balanced, eclectic view in the 1990s versus a skill orientation in the 1960s), instructional organization (more whole-class instruction in the 1990s vs. ability grouping in the 1960s), instructional materials (more use of trade books in the 1990s vs. primarily basal readers in the 1960s), early literacy instruction (the reading readiness perspective of the 1960s was supplanted by an emergent literacy orientation), and assessment (use of more nonstandardized, alternate assessments in the 1990s).

The authors noted that there were several parallel concerns expressed by teachers of the 1960s and 1990s—most significantly, an urgency to accommodate struggling readers—but they also commented that contemporary teachers and administrators were more likely to modify programs and explore alternate philosophical orientations than teachers and administrators of the past. They noted that reading instruction of the 1990s was "not some gussied-up version of the good old days," but instead they found "great energy," "a commitment to children, teaching, and learning," and "a sense of motivated urgency to adopt instructional principles, practices, and philosophies that will accommodate learners of today" (Baumann, Hoffman, et al., 2000, p. 361).

SUMMARY

The purpose of this chapter was to provide an overview of survey research methods employed in literacy research. Survey research is a method for gathering information from a sample of individuals in order to describe characteristics of a larger group. Survey research has been employed in education and social sciences for the past 100 years, with increasing activity in the past three decades. Commonly used survey methods involve collecting data through face-to-face, telephone, and mail surveys, with recent increase in the use of technology-enhanced methods. Most surveys involve probability samples from a specified population in order to make generalizations about the broader population. There is a generally accepted and implemented survey research design process that leads researchers from question and hypothesis formulation through sample description and data collection to data analyses and survey reporting. Quality standards can be applied to evaluate published surveys and to guide re-

searchers in designing survey studies. The McKenna et al. (1995) survey of elementary children's reading attitudes and the Baumann, Hoffman, et al. (2000) survey of elementary reading instruction practices were presented as exemplars of sound literacy education survey research.

REFERENCES

Austin, M. C., Bush, C. L., & Huebner, M. H. (1961). *Reading evaluation: Appraisal techniques for school and classroom.* New York: Ronald Press.

Austin, M. C., & Morrison, C., with Kenney, H. J., Morrison, M. B., Gutmann, A. R., & Nystrom, J. W. (1961). *The torch lighters: Tomorrow's teachers of reading.* Cambridge, MA: Harvard Graduate School of Education/Harvard University Press.

Austin, M. C., & Morrison, C., with Morrison, M. B., Sipay, E. R., Gutmann, A. R., Torrant, K. E., & Woodbury, C. A. (1963). *The first R: The Harvard report on reading in elementary schools.* New York: Macmillan.

Barton, A. H., & Wilder, D. E. (1964). Research and practice in the teaching of reading: A progress report. In M. B. Miles (Ed.), *Innovation in education* (pp. 361–398). New York: Teachers College, Columbia University.

Baumann, J. F., Hoffman, J. V., Duffy-Hester, A. M., & Ro, J. M. (2000). *The first R* yesterday and today: U.S. elementary reading instruction practices reported by teachers and administrators. *Reading Research Quarterly, 35,* 338–377.

Baumann, J. F., Ro, J. M., Duffy-Hester, A. M., & Hoffman, J. V. (2000). Then and now: Perspectives on the status of elementary reading instruction by prominent reading educators. *Reading Research and Instruction, 39,* 236–264.

Bloor, M., Frankland, J., Thomas, M., & Robson, K. (2001). *Focus groups in social research.* Thousand Oaks, CA: Sage.

Commeyras, M., & DeGroff, L. (1998). Literacy professionals' perspectives on professional development and pedagogy: A United States survey. *Reading Research Quarterly, 33,* 434–472.

Converse, P. E. (1964). The nature of belief systems in mass publics. In D. E. Apter (Ed.), *Ideology and discontent* (pp. 207–261). New York: Free Press.

Couper, M. P. (2002). Web surveys: A review of issues and approaches. *Public Opinion Quarterly, 64,* 464–494.

Couper, M. P., Baker, R. P., Bethlehem, J., Clark, C. Z. F., Martin, J., Nicholls, W. L., & O'Reilly, J. M. (1998). *Computer assisted survey information collection.* New York: Wiley.

Dillman, D. A. (1978). *Mail and telephone surveys: The total design method.* New York: Wiley.

Flippo, R. F. (1998). Points of agreement: A display of professional unity in our field. *The Reading Teacher, 52,* 30–40.

Fowler, F. J. (2002). *Survey research methods* (3rd ed.). Thousand Oaks, CA: Sage.

Greenbaum, T. L. (1998). *The handbook of focus group research*. Thousand Oaks, CA: Sage.

Groves, R. M., Dillman, D. A., Eltinge, J. L., & Little, R. J. A. (Eds.). (2002) *Survey nonresponse*. New York: Wiley.

Gubrium, J. F., & Holstein, J. A. (Eds.). (2002). *Handbook of interview research: Context and method*. Thousand Oaks, CA: Sage.

Hochstim, J. R. (1967). A critical comparison of three strategies of collecting data from households. *Journal of the American Statistical Association, 62,* 976–989.

Hoffman, J. M., Assaf, L. C., & Paris, S. G. (2001). High-stakes testing in reading: Today in Texas, tomorrow? *The Reading Teacher, 54,* 482–492.

Hoffman, J. V., & Roller, C. M. (2001). The IRA Excellence in Reading Teacher Preparation Commission's report: Current practices in reading teacher education at the undergraduate level in the United States. In C. M. Roller (Ed.), *Learning to teach reading: Setting the research agenda* (pp. 32–79). Newark, DE: International Reading Association.

Hughes, M. T., Cash, M. M., Klingner, J., & Ahwee, S. (2001). Professional development programs in reading: A national survey of district directors. In J. V. Hoffman, D. L. Schallert, C. M. Fairbanks, J. Worthy, & B. Maloch (Eds.), *Fiftieth yearbook of the National Reading Conference* (pp. 275–286). Chicago: National Reading Conference.

Ivey, G., & Broaddus, K. (2001). "Just plain reading": A survey of what makes students want to read in middle school classrooms. *Reading Research Quarterly, 36,* 350–377.

Jaeger, R. M. (1997). Survey research methods in education. In R. M. Jaeger (Ed.), *Complementary methods for research in education* (2nd ed., pp. 449–476). Washington, DC: American Educational Research Association.

Kalton, G. (1983). *Introduction to survey sampling*. Beverly Hills, CA: Sage.

Kraemer, H. C., & Thiemann, S. (1987). *How many subjects? Statistical power analysis in research*. Thousand Oaks, CA: Sage.

Krosnick, J. A. (1999). Survey research. *Annual Review of Psychology, 50,* 537–567.

Krueger, R. A., & Casey, M. A. (2000). *Focus groups: A practical guide for applied research* (3rd ed.). Thousand Oaks, CA: Sage.

Link, M. W., & Oldendick, R. W. (1999). Call screening: Is it really a problem for survey research? *Public Opinion Quarterly, 63,* 577–589.

Matthewson, G. C. (1994). Model of attitude influence upon reading and learning to read. In R. B. Ruddell, M. R. Ruddell, & H. Singer (Eds.), *Theoretical models and processes of reading* (4th ed., pp. 1131–1161). Newark, DE: International Reading Association.

McKenna, M. C. (1994). Toward a model of reading attitude acquisition. In E. H. Cramer & M. Castle (Eds.), *Fostering the life-long love of reading: The affective domain in reading education* (pp. 18–40). Newark, DE: International Reading Association.

McKenna, M. C., & Kear, J. D. (1990). Measuring attitude toward reading: A new tool for teachers. *The Reading Teacher, 43,* 626–639.

McKenna, M. C., Kear, D. J., & Ellsworth, R. A. (1995). Children's attitudes

toward reading: A national survey. *Reading Research Quarterly, 30,* 934–956.

More about mail surveys. (1997). American Statistical Association brochures about survey research. Retrieved online February 6, 2003, from *http://www.amstat.org/sections/srms/brochures/Mail.pdf*

Morrison, C. (1963). *A critical analysis of reported and recommended reading practices in the elementary schools.* Unpublished doctoral dissertation, Harvard Graduate School of Education, Cambridge, MA.

Morrison, C., & Austin, M. C. (1977). *The torch lighters revisited.* Newark, DE: International Reading Association.

Nachmias, D., & Nachmias, C. (1981). *Research methods in the social sciences.* New York: St. Martin's Press.

Publication manual of the American Psychological Association (5th ed.). (2001). Washington, DC: American Psychological Association.

Rose, L. C., & Gallup, A. M. (2002). The 34th annual Phi Delta Kappa/Gallup poll of the public's attitudes toward the public schools. *Phi Delta Kappan, 84*(1), 41–56.

Schaefer, D. R., & Dillman, D. A. (1998). Development of a standard e-mail methodology: Results of an experiment. *Public Opinion Quarterly, 62,* 378–397.

Seidman, I. (1998). *Interviewing as qualitative research: A guide for researchers in education and the social sciences* (2nd ed.). New York: Teachers College Press.

Shulman, L. S. (1997). Disciplines of inquiry in education: A new overview. In R. M. Jaeger (Ed.), *Complementary methods for research in education* (2nd ed., pp. 3–29). Washington, DC: American Educational Research Association.

Strunk, W., & White, E. B. (2000). *The elements of style* (4th ed.). Boston: Allyn & Bacon.

Tucker, C., Lepkowski, J. M., & Piekarski, L. (2002). List-assisted telephone sampling design efficiency. *Public Opinion Quarterly, 66,* 321–338.

Venezky, R. L. (1984). The history of reading research. In R. Barr, M. L. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (pp. 3–38). New York: Longman.

Weisberg, H. F., Krosnick, J. A., & Bowen, B. D. (1996). *An introduction to survey research, polling, and data analysis.* Thousand Oaks, CA: Sage.

What are focus groups? (1997). American Statistical Association brochures about survey research. Retrieved online February 6, 2003, from *http://www.amstat.org/sections/srms/brochures/focusgroups.pdf*

What is a margin of error? (1998). American Statistical Association brochures about survey research. Retrieved online February 6, 2003, from *http://www.amstat.org/sections/srms/brochures/margin.pdf*

What is a survey? (n.d.). American Statistical Association brochures about survey research. Retrieved online February 6, 2003, from *http://www.amstat.org/sections/srms/brochures/survwhat.html*

Wiseman, F. (1972). Methodological bias in public opinion surveys. *Public Opinion Quarterly, 36,* 105–108.