

Mass Communication Research and Theory

EDITED BY

Guido H. Stempel III

Ohio University

David H. Weaver

Indiana University

G. Cleveland Wilhoit

Indiana University

PUB 2003



*Boston • New York • San Francisco
Mexico City • Montreal • Toronto • London • Madrid • Munich • Paris
Hong Kong • Singapore • Tokyo • Cape Town • Sydney*

Survey Research

Pamela J. Shoemaker

Maxwell E. McCombs

Media professionals spend a lot of time dealing with the products of survey research. Journalists write public opinion poll stories to tell Americans who's going to win the next election or to report what we think about the latest issue. Advertisers assure us that "9 out of every 10 dentists surveyed" support their products. And media management's marketing strategies are based on what surveys say the audience wants. Surveys have even invaded the Internet.

A *survey* is a study that collects information by asking people questions. The information collected—the data—is generally numerical and suitable for statistical analysis. Although the United States Constitution specified that a survey of Americans be conducted every ten years in order to ensure proper representation in the House of Representatives,¹ this census of all U.S. households is not what we commonly think of as a survey. The vast majority of survey research projects are *sample surveys* in which data are collected from a subset of individuals in the population. Inferences about the larger population are made from the information gathered from those people in the sample. U.S. public opinion polls commonly include survey interviews with 1,000 to 1,200 people, even though the adult population of the United States exceeds that by millions.

Surveys have proven to be an effective way to assess people's voting intentions, to explain their votes, to predict use of products, and to assess changes in opinions. In fact, the effectiveness of the telephone survey has resulted in the public being bombarded with polls. Unfortunately, many polls are thinly disguised attempts to sell products and services, thus making survey research more difficult for the legitimate pollster.

Surveys of Americans' actual or intended votes have been conducted for 200 years, with one of the earliest polls on record being conducted after the election for Massachusetts governor in 1787.² Supporters of the incumbent Governor James Bowdoin conducted a poll to explain why John Hancock won the race. The Bowdoin poll showed that Hancock's support came primarily from "labourers, servants, and so on," whereas Bowdoin's base was among "merchants and traders."

Three days later a Hancock supporter revealed his own poll, a recategorization of the vote that he called the “authentic breakdown,” which showed that Bowdoin’s support was actually among “speculators in public securities,” whereas Hancock’s vote came primarily from “merchants, tradesmen, and other worthy members of society” and from “friends to the revolution.” The Hancock poll also revealed that the one wizard voting in the election supported the losing candidate!³

Although such canvassing of voters became fairly common in the 1800s, surveys of people’s opinions on issues of the day were rare. One of the first U.S. opinion polls was conducted in the 1880s by the Iowa Labor Statistics Bureau. As part of its survey on labor and economic conditions, the bureau asked farmers’ attitudes on the liquor issue and immigration.⁴

The advent of modern random sampling techniques in the early 1900s made surveys more practical, since a smaller number of interviews meant savings in time and expense,⁵ but surveys using random samples also proved to be more accurate. In the 1936 U.S. presidential election, pollsters George Gallup, Elmo Roper, and Archibald Crossley used the expertise of statisticians, psychologists, and market researchers to correctly predict Franklin D. Roosevelt’s victory over Alf Landon, whereas the less methodologically sound *Literary Digest* magazine poll of 10 million people picked Landon by a wide margin.⁶

Gallup and other pollsters enjoyed popularity until they predicted incorrectly that Thomas Dewey would beat Harry Truman in 1948. Confidence in polling collapsed until in the 1960s, when secret polls captured the imagination of politicians.⁷ John Kennedy used polls in his 1960 election campaign, thus dignifying them for use by others, and Jimmy Carter was the first U.S. president to have access to continuous opinion polls.⁸

Much of the interviewing in opinion polls now is being done by telephone, rather than in face-to-face interviews, because of the rising costs and lower response rates face-to-face interviewing has yielded in recent years.⁹ About 98 percent of U.S. households are reachable by telephone.¹⁰ Internet polling is also becoming increasingly popular.

Survey Designs and Applications

One of the first decisions to be made in planning a survey involves the *design*—who will be interviewed, over what period, and how many times? There are three basic survey designs: the **cross-sectional** survey, the **panel study**, and the **trend study**.

The decision about whom to interview is largely a theoretical one driven by the hypotheses that are to be tested, but it is not a trivial decision. For example, in an election survey, we could interview adults, or only adult citizens, or only registered voters, or only people who actually intend to vote. The population we select for study will affect the data we collect and the conclusions we draw.

The period of time during which interviews are conducted—generally called being “in the field”—is also crucial. Some polls are in the field for less than one day, meaning that interviews have been conducted only with those individuals who are home at the instant that the interviewer calls. Such a sample of respondents may underrepresent busy individuals. Other polls are conducted over several months, making the results vulnerable to events occurring while the study is in the field. For example, a poll asking about people’s

attitudes toward AIDS could yield two very different data sets if a cure for AIDS is revealed in the middle of the field period. The optimal field period is long enough to make several attempts to complete an interview in each sampled household, but short enough to minimize threats to validity due to intervening events.

In a *cross-sectional* survey, the respondents are interviewed only once, and the data collected provide a snapshot of the population at the time the field work (the interviewing) is done. Cross-sectional surveys are common, because they are far more economical than longitudinal studies. Cross-sectional surveys can describe the characteristics of the sample (e.g., respondents' attitudes or behaviors) at a given point in time, but such one-shot studies cannot address questions about change (e.g., is a change in a person's attitude followed by a change in his or her behavior?). To study changes among the elements in a sample (and, by inference, in the population from which the sample is drawn) requires that data be collected at more than one point in time—a *longitudinal* study. The panel study and the trend study both involve longitudinal data collection.

The *panel study* involves interviews with the same individuals at more than one point in time. The panel design permits the comparison of a person's responses at each time point, thereby allowing us to assess the extent and direction of any changes that may have occurred *within that individual*. For example, we might compare changes in an individual's media use with changes in support for a political candidate.

The *trend study* is also used to study change, but change in the population as a whole rather than changes in individuals. In a trend study, a new sample of individuals is drawn from the population at each point in time—the trend study is composed of several cross-sectional studies that are compared over time. This permits the researcher to evaluate change in the population over time (e.g., changes in support for two presidential candidates); however, because such changes (e.g., in vote intention) are attributes of individuals, the trend design limits our ability to explain change. Because different individuals are interviewed at each point in time, we do not know why the observed changes have occurred, and we cannot easily separate changes in individuals over time from possible differences between the samples drawn at each time point.

The panel is a superior design for explaining change, but data from a panel study may be flawed by *panel mortality*—the loss of individuals from the panel between each interviewing wave. Individuals leave panels for many reasons: They decline to participate; they move during the study or change telephone numbers and are now unreachable; they go on vacation or business trips or become ill; or they may even die during the study. Panel mortality is generally not the result of a random process, and therefore the individuals who leave a panel generally have characteristics in common. Their failure to participate in the entire study may make the panel unrepresentative of the population, thereby limiting the inferences we can make about the population.

Attempts to limit panel mortality can make the panel study more expensive than the trend design, but the panel remains the more valid way of assessing changes in individuals. The trend design is useful in the *secondary analysis* of preexisting data sets. Many survey organizations (such as the University of Michigan Institute for Social Research and the National Opinion Research Center) archive survey data and make data sets available to others. Secondary analysis is "the extraction of knowledge on topics other than those which were the focus of the original surveys."¹¹ When the same questions are asked in

successive surveys, the secondary analyst can perform a trend analysis of several data sets. For example, McCombs used the well-known Michigan national election surveys to trace increasing political participation by blacks through their use of newspapers and television for political information during the 1952–1964 civil rights era.¹²

A fourth design is not recommended—the *pseudopanel*. If a researcher asks her respondents to report their past and current attitudes or behaviors, she may decide to treat her data as if they represent different time points. Such a design is vulnerable to many problems, however, not the least of which is the human being's ability or willingness to remember and report what he or she thought or did in the past. The pseudopanel design is subject to many potential errors and should be avoided.

Using Survey Data to Establish Causal Relationships

Establishing a causal relationship requires: (1) showing covariation between the presumed cause and the presumed effect; (2) establishing appropriate time order, that is, that the cause precedes the effect and not the reverse; (3) ruling out alternative explanations for the observed relationship; and (4) controlling error variance, that is, keeping errors to a minimum and explaining as much of the independent variable as possible.

The cross-sectional design—representing the bulk of all survey projects—can address only three of the four criteria necessary for establishing a causal relationship: covariation, alternative explanations, and error. Covariation can be shown between two or more variables through the use of statistics that test the strength of the relationship between two variables (e.g., Pearson's correlation coefficient). But, as the saying goes, “correlation is not causation”—showing that changes in one variable (e.g., exposure to a daily newspaper) are associated with changes in another (e.g., levels of political interest) is only one step, but a necessary one, toward proving causality.

Survey researchers have become adept at *multivariate statistical techniques* that help rule out alternative explanations for a relationship by statistically controlling for the effects of other variables on the observed relationship. The use of statistical controls, however, is only as good as the researcher's ability to identify and measure possible alternative explanations. It is far easier to rule out alternative explanations in a randomized experiment, where the experimenter randomly assigns subjects to treatment groups, thereby controlling for a vast array of unidentified and unmeasured variables. In a survey, respondents select themselves into high or low political interest groups, thus making differences in these groups related to many causes.

Time order can be established in survey research through the use of longitudinal designs, with the panel's ability to study changes within individuals making it far superior to the trend or cross-sectional design in establishing causality. In a panel study, we can look at whether, for example, individuals first change their level of newspaper reading and then their level of political interest, or vice versa. If we observe that newspaper reading and political interest change simultaneously, then we have evidence of a third variable affecting both reading and interest or a methodological problem with the time intervals between our panel waves. The waves of interviewing need to be frequent enough to “catch” changes in variables that may occur rapidly one after the other.