

Pearson New International Edition

Research in Education

John W. Best James V. Kahn Tenth Edition

From Chapter 5 of *Research in Education*, Tenth Edition. John W. Best, James V. Kahn. Copyright © 2006 by Pearson Education, Inc. All rights reserved.

Assessment, Evaluation, and Research

A descriptive study describes and interprets what *is*. It is concerned with conditions or relationships that exist, opinions that are held, processes that are going on, effects that are evident, or trends that are developing. It is primarily concerned with the present, although it often considers past events and influences as they relate to current conditions.

The term *descriptive study* masks an important distinction, for not all descriptive studies fall into the category of research. In fact, of the three kinds of descriptive studies included in this chapter, only one is actually research. Also, this chapter confines itself to descriptive studies that use quantitative methods.

Assessment describes the status of a phenomenon at a particular time. It describes without value judgment a prevailing situation; it attempts no explanation of underlying reasons and makes no recommendations for action. It may deal with prevailing opinion, knowledge, practices, or conditions. As it is ordinarily used in education, assessment describes the progress students have made toward educational goals at a particular time. For example, in the National Assessment of Education Progress program, the data are gathered by a testing program and a sampling procedure in such a way that no individual is tested over the entire test battery. It is not designed to determine the effectiveness of a particular process or program but merely to estimate the degree of achievement of a large number of individuals exposed to a great variety of educational and environmental influences. It does not generally provide recommendations, but there may be some implied judgment on the satisfactoriness of the situation or the fulfillment of society's expectations.

Evaluation is a process used to determine what has happened during a given activity or in an institution. The purpose of evaluation is to see if a given program is working, if an institution is successful according to the goals set for it, or if the original intent is being successfully carried out. To assessment, evaluation adds the ingredient of value judgment of the social utility, desirability, or effectiveness of a process, product, or program, and it sometimes includes a recommendation for some course of action. School surveys are usually evaluation studies; educational products and programs are examined to determine their effectiveness in meeting accepted objectives, often with recommendations for constructive action.

Descriptive research deals with the relationships between variables, the testing of hypotheses, and the development of generalizations, principles, or theories that have universal validity. The expectation is that, if variable A is systematically associated with variable B, prediction of future phenomena may be possible, and the results may suggest additional or competing hypotheses to test. Descriptive research is sometimes divided into correlational research, causal–comparative research, and other descriptive research that is neither correlational nor designed to find causation but describes existing conditions. All of these types of descriptive research are included here because they have the same basic components: They are all attempting to find generalizable attributes, and they all deal with present conditions.

In carrying out a descriptive research project, in contrast to an experiment, the researcher does not manipulate the variable, decide who receives the treatment, or arrange for events to happen. In fact, the events that are observed and described would have happened even if there had been no observation or analysis. Descriptive research also involves events that have already taken place and may be related to a present condition.

The method of descriptive research is particularly appropriate in the behavioral sciences because many of the types of behavior that interest the researcher cannot be arranged in a realistic setting. Introducing significant variables may be harmful or threatening to human subjects. Ethical considerations often preclude exposing human subjects to harmful manipulation. For example, it would be unthinkable for an experimenter to randomly decide who should smoke cigarettes and who should not smoke them for the purpose of studying the effect of smoking on cancer, heart disease, or other illnesses thought to be caused by cigarette smoke. Similarly, to deliberately arrange auto accidents, except when mannequins are used, in order to evaluate the effectiveness of seat belts or other restraints in preventing serious injury would be absurd.

Although many experimental studies of human behavior can be appropriately carried out both in the laboratory and in the field, the prevailing research method of the behavioral sciences is descriptive. Under the conditions that naturally occur in the home, the classroom, the recreational center, the office, or the factory, human behavior can be systematically examined and analyzed.

The many similarities between these types of descriptive studies may cloud the distinctions between them. They are all characterized by disciplined inquiry, which requires expertise, objectivity, and careful execution. They all develop knowledge, adding to what is already known. They use similar techniques of observation,

description, and analysis. The differences between them lie in the motivation of the investigator, the treatment of the data, the nature of the possible conclusions, and the use of the findings. The critical distinctions are that the three types of studies have different purposes and, therefore, approach the problem differently and that only descriptive research studies lead to generalizations beyond the given sample and situation.

A single study may also have multiple purposes. For instance, a study may evaluate the success or failure of an innovative program and also include sufficient controls to qualify as a descriptive research study. Similarly, an assessment study may include elements that result in descriptive research also. Unfortunately, this potential overlap further clouds the distinction. Put simply, in order for a descriptive study to be considered research, it must have sufficient controls to permit generalization of the results.

Examples of these three types of descriptive studies follow. It is important to keep in mind that, although these examples are presented to illustrate each individual type of study (assessment, evaluation, or descriptive research), they are not mutually exclusive. That is, for example, although surveys are used to illustrate assessment and evaluation studies, surveys are also used in descriptive research studies. Similarly, although causal-comparative studies illustrate their major use in descriptive research, this type of design can also be used in an assessment or evaluation study.

Assessment Studies

Since the publication of A Nation at Risk (National Commission on Excellence in Education, 1983), there has been an increased call for greater accountability of the public schools. The result is educational standards and/or curriculum guides in every one of the 50 states. As a result of the federal No Child Left Behind (NCLB) Act of 2001, all states require tests to determine if both the standards are achieved by individual students and the percentage of "success" at given schools and school districts. In addition, there are many persons (particularly politicians) who would like to see national standards coupled with a test.

No Child Left Behind

The No Child Left Behind Act has turned into a controversial piece of legislation with many school districts complaining about unrealistic expectations and high Navigator.com costs. The reader is referred to the Research Navigator search for more information.

As part of school reform in Chicago, students in the Chicago Public Schools who do not pass a test at the end of the 8th grade must go to summer school and/or repeat the 8th grade. Schools are given "report cards" and can be put on probationary status or "reconstituted." When schools are reconstituted, the principal and teachers are no longer guaranteed positions at that school. The school hires new personnel (which may include some of the teachers who had already been working there) under the supervision of the central office staff. In other words, an attempt is made to start over with a "new school."

Many parents and teachers feel that the proliferation of testing takes too much time (both the time for taking the test and the time many teachers spend preparing their students for the test) away from instructional time. Also, the greater emphasis on tests creates more pressure on students. While almost everyone agrees with the need for accountability, there are differing views regarding how best to achieve it.

Some of the assessments conducted under the guise of accountability are actual assessment studies. These studies usually fall into the categories of surveys and educational assessment data. These latter data are generally used to compare school districts within the state that collected the data and to compare individual schools within the districts. Data regarding individual students that are not aggregated should not be considered part of an assessment study. Assessment studies include various types of surveys, educational assessments, activity analysis, and trend studies.



ator.com

Survey Research

The Survey

The survey method gathers data from a relatively large number of cases at a particular time. It is not concerned with characteristics of individuals *as* individuals. It is concerned with the statistics that result when data are abstracted from a number of individual cases. It is essentially cross-sectional.

Ninety-four percent of American homes have at least one television set. About three out of five students who enter American secondary schools remain to graduate. Seventy-one percent of Americans say they favor a constitutional amendment to permit school prayer, whereas only 45% would still favor school prayer if it offended a large percentage of parents. In 1990, 4,717,641 children were enrolled in California public schools, whereas slightly more than 10% of that number, 474,194, were enrolled in Iowa's public schools. On an average day, 11,137 babies are born. In both 1987 and 1995, 77% of those polled favored an increase in the minimum wage. The population of Illinois according to the 1980 census was 11,426,518. The Illinois population increased slightly according to the 1990 census figures to 11,430,602. Data like these result from many types of surveys. Each statement pictures a prevailing condition at a particular time. All of the survey data above were obtained from resources on the Internet, which provides a new method of disseminating information much more quickly than through the usual publication channels.

In analyzing political, social, or economic conditions, one of the first steps is to get the facts about the situation or a picture of conditions that prevail or that are developing. These data may be gathered from surveys of the entire population. Others are inferred from a study of a sample group carefully selected from the total population. At times the survey may describe a limited population that is the only group under consideration.

The survey is an important type of study. It must not be confused with the mere clerical routine of gathering and tabulating figures. It involves a clearly defined problem and definite objectives. It requires expert and imaginative planning, careful analysis and interpretation of the data gathered, and logical and skillful reporting of the findings.

Social Surveys

Surveys became a common method of data collection in the first half of this century. In the late 1930s a significant social survey was directed by the Swedish sociologist Gunnar Myrdal and sponsored by the Carnegie Foundation. Myrdal and his staff of researchers made a comprehensive analysis of the social, political, and economic life of African Americans in the United States, yielding a great mass of data on race relations in America (Myrdal, 1944).

The late Alfred Kinsey of Indiana University made a comprehensive survey of the sexual behavior of the human male, based on data gathered from more than 12,000 cases (Kinsey, Pomeroy, & Martin, 1948). His second study (Kinsey, Pomeroy, Martin, & Gebhard, 1953) of the behavior of the human female followed later. Studies of sexual behavior have continued to the present. However, we should note that there are problems with surveys of sensitive topics such as these (Zimmerman & Langer, 1995). The most obvious of the problems with surveys of sensitive behaviors is whether the responses are honest. Because some people may exaggerate and others understate their behaviors, the data from such surveys are often hard to interpret. Zimmerman and Langer suggest ways to optimize the estimates regarding the rates of sensitive behaviors. Although studies on sexual behaviors have raised considerable controversy, they represent a scientific approach to the study of an important social problem and have many implications for jurists, legislators, social workers, and educators.

Witty (1967) studied the television viewing habits of school children and published annual reports on his investigations. These studies were conducted in the Chicago area and indicate the amount of time devoted to viewing and the program preferences of elementary and secondary students, their parents, and their teachers. Witty attempted to relate television viewing to intelligence, reading habits, academic achievement, and other factors. Recent survey data (July, 1996) on television viewing reported by the American Academy of Pediatrics (http://edie.cprost.sfu.ca/gcnet/iss4-21c.html) indicate that American children and adolescents spend between 22 and 28 hours a week watching television, more than any other activity except sleeping. The American Psychological Association estimates that the average American child or teenager views 10,000 murders, rapes, and aggravated assaults on television each year. More than 1,000 studies have indicated that exposure to heavy doses of television violence is highly related to the likelihood of aggressive behavior, particularly in males. Unfortunately, these data and studies of aggressive behavior have not led to changes in the television programming or viewer behavior despite the fact that almost 80% of Americans believe that violence on television contributes to violence in society.

Shaw and McKay (1942) conducted a study of juvenile delinquency in Chicago, yielding significant data on the nature and extent of delinquency in large urban communities. Lang and Kahn (1986) examined special education teacher estimates of their students' criminal acts and crime victimizations. The data indicated that special education students seem to be victimized in the same way as others but to a greater degree. This preliminary study led to Lang's (1987) dis-