Stanfield, John. 2011. Rethinking Race and Ethnicity in Research Methods. Walnut Creek, CA: Left Coast Press.

7 Researching Race and Ethnicity: (Re)Thinking Experiments

Henry A. Walker

> Theory is the method of the sciences.
>
> —David Willer (1987:ix)

Introduction

This chapter is concerned with the experiment as a technique for investigating phenomena in the field of race and ethnic studies. My view is that a discussion of experiments, or of any other research method, is incomplete unless it is embedded in a discussion of the role that research plays in the scientific process.[1] Science is concerned with developing and testing explanations for relationships between phenomena (i.e., theories). Mature sciences like chemistry and physics are marked by two characteristics: (1) they have well-developed bodies of theory; and (2) they make extensive use of experiments. In contrast, theory is underdeveloped in the study of race and ethnicity, and social science experiments on race matters are statistically rare.

Disciplinary differences in theory development and in the incidence of experimental research are quantitative not qualitative but they reflect a deeper problem for the social sciences. Theory is *the* method of science; it distinguishes the sciences from other academic disciplines. The degree to which theory is central to the research process differs in the physical and social sciences. Physical and social scientists test hypotheses but, unlike their physical science counterparts, social scientists typically derive their hypotheses from ad hoc speculation and prior observations rather than theory. It is not coincidental that experimental studies of race are rare and that theory development in the study of race and ethnicity lags behind theory development in fields that make greater use of experimentation.

By and large, the mature sciences and the social sciences take different approaches to science. An *empiricist* approach to science dominates work in sociology and race and ethnic studies. The approach taken in the mature

sciences is more often *theory driven*. In this chapter, I argue that widespread adoption of the theory-driven approach is necessary to transform race and ethnic studies from a field that amasses unexplained empirical observations to a science that offers more, and better, understandings of human behavior. Experiments are an important element of the approach. I also assert that combining a theory-driven approach with the norms of scientific conduct can reduce (if not eliminate) many of the problems that critics ascribe to research methods. The time is ripe for social scientists to reconsider an investigative tool to which most have received only a cursory introduction. But we must understand the nature of experiments and their role in the scientific process before we can seriously *rethink* their use. I came to this position easily, but the path may be more difficult for some.

I have designed and run many sociological experiments during a career that spans almost four decades. Before embarking on a career in sociology, I spent countless hours running biology and chemistry experiments in high school and college laboratories. My earliest exposure to social science research methods came in courses that fulfilled requirements for an undergraduate degree in sociology. Each methods course included material on experiments and there was much talk of theory in those courses. But there was little actual theory. Sociology introduced me to the empiricist approach to science.[2]

I got my social science introduction to theoretical methods as a graduate student at the University of Missouri at Kansas City. The late Ernest Manheim[3] required students to read Hempel and Oppenheim (1948), Nagel (1961), Kaplan (1964), and others. At Stanford, I got more exposure in course work with Joseph Berger, Bernard P. Cohen, and Morris Zelditch, Jr. Collectively, those instructors reintroduced me to the theory-driven approach to science and its connection to experimentation. Given my earlier experiences as a student in the physical sciences, I had no difficulty adopting an approach that conceived of science as a continuous process of: (1) theory development; (2) theory testing; and (3) theory refinement.

To make the case for a theory-driven approach, the remainder of this chapter is organized as follows: In the next section, I discuss problems associated with the conduct of inquiry in race studies, including the statistical rarity of experiments as an understudied phenomenon. I discuss empiricist and theory-driven approaches to science in the third section. There, I argue that adopting a theory-driven strategy is crucial to advancing science. The fourth section is devoted to a discussion of data collection and reasons that scientists collect data. The discussion locates research and research methods in the larger scientific enterprise and corrects common misunderstandings about science, experiments, and experimental research. In the fifth section, I give examples of experiments to show how they have been used (and misused) to study race and ethnic issues. The last section includes an appeal for more

theory-driven research and more extensive use of experiments in studies of race and ethnicity.

Sociological Studies of Race and Ethnicity: Dilemmas and Problems

The field of race and ethnic relations encompasses a broad range of phenomena. At the microsocial level, investigators are concerned with patterns of conscious and unconscious bias (Greenwald et al. 1998), the relationship between individuals' prejudices and discrimination (LaPiere 1934), and how discrimination affects the well-being of its targets (Allport 1954). Macrosocial researchers study societal patterns of ethnic and race relations such as lynchings in the nineteenth-century United States (Wells-Barnett [1892–1895/2002]), global patterns of slavery (Patterson 1982), and race and class stratification in modern America (Wilson 2009). Between the micro and macro extremes, students of race and ethnic relations study every imaginable human behavior, including crime and punishment (Mann 1993), academic achievement (Ogbu 2003), patterns of family formation and family structure (Moynihan 1965), and preferences for the race composition of neighborhoods (Clark 2002; Williams 1975).

Social scientists engaged in the systematic study of race and ethnicity have a wide array of research techniques at their disposal. That is as it should be in a field that encompasses such a broad range of phenomena. The new field began to blossom at the turn of the twentieth century, as did criticisms of researchers and the methods they used to study race. Minority scholars were among the earliest contributors to the critical literature.[4] Du Bois proclaimed more than a century ago: "[Sociologists] simply collect the facts. Others may use them as they will" (cited in Green & Driver 1976:313).

Du Bois understood that collecting facts is not a simple matter and that interpreting and applying facts to problems of the human condition are difficult at best. As a septuagenarian reflecting on his youthful search for scientific truth, Du Bois pointed out that "[At the time] the difficulties of applying scientific law and discovering cause and effect in the social world were still great" (Du Bois 1940:50–51). His concerns were not limited to questions of theory development (i.e., discovering cause and effect) or the application of theory. Nor did he claim that those problems—as important as they are—were the most important barriers to developing knowledge about race matters. He said:

> Most unfortunate of all, however, is the fact that so much of the work done on the Negro question is notoriously uncritical; uncritical from lack of discrimination in the selection and weighing of evidence; uncritical in choosing the proper point of view from which to study these problems, and,

finally, uncritical from the distinct *bias* in the minds of so many writers. (Du Bois 1898:12–13; emphasis added)

Du Bois's concerns presaged unfavorable evaluations offered by contemporary scholars who have studied the history of research on race and ethnicity—fact collection—and found it wanting. Mainstream sociology is taken to task on epistemological grounds (Ladner 1973:xx), with some labeling its logical positivist foundation a flawed or limited approach (Stanfield & Dennis 1993:16ff.). Critics also raise ontological concerns about the meaning of race as a biological construct (Graves 2004; Montagu 1942) and question the use of specific research techniques. (See papers on survey research, community studies, participant observation, demographic analyses of population data, etc., in Ladner [1973], Stanfield & Dennis [1993], and Zuberi & Bonilla-Silva [2008].)

Modern-day criticisms center on additional issues that include but are not limited to: (1) underrepresentation of minority scholars and minority subjects in sociological research; (2) cultural, political, and ideological biases that affect researchers' observations, their interpretations of research, and the explanations they offer for race and ethnic phenomena; (3) unethical behavior, including maltreatment of minority subjects and their communities; (4) researchers' positive evaluations of the status quo (i.e., majority dominance) vis-à-vis alternative ways of organizing social life; and (5) the difficulty of translating sociological knowledge into policies that secure *positive* results. With increasing frequency, contemporary critics express skepticism concerning the capacity of sociology to uncover social facts and to develop valid explanations for them. Many reputable scholars also question the ability of sociological research to inform policies that can advance the human condition (Cole 1994).

Experimental studies of race have generally escaped the criticisms directed toward other research techniques. Goar (2008) calls attention to the oversight and points out that many early experiments on race and ethnicity are contaminated by bias. Moreover, she asserts that white privilege has sustained the pattern in subsequent experimental research. Importantly, Goar uncovers reasons for the scarcity of critical commentary on experimental methods. She points to evidence from Hunt et al. (2000) that race and ethnicity are understudied topics in social psychology. Social psychologists use experiments more often than any other sociologists and *Social Psychology Quarterly* (SPQ) is the leading journal of research in sociological social psychology. Hunt and colleagues analyzed 954 articles published in SPQ from 1970 to 1999 and reported that only 8.3% "seriously considered" race and ethnicity. Unfortunately, Hunt and colleagues did not report data on research techniques.

To correct this omission, I reviewed all articles published in four leading journals (*American Journal of Sociology* [AJS], *American Sociological Review* [ASR], *Social Forces* [SF], and *Social Psychology Quarterly* [SPQ]) and two specialty journals (*Ethnic & Racial Studies* [ERS] and *Social Problems* [SP]) that are more likely to include studies of race. I counted 482 articles, including eighty-eight in ERS, published in calendar years 2007 and 2008.[5] Using less stringent inclusion criteria than Hunt and colleagues, I found that 180 articles (37.3% of the total) focused on race or used race (ethnicity) as an explanatory variable. Omitting articles from ERS, 107 of the remaining 394 articles (27.2%) focus on race. Fifteen articles (3.1% of 482) reported using the experimental method, but only two race studies (.4% of all studies and 1.1% of 180 race studies) reported using experiments. SF and ERS each published a single race experiment. Despite publishing only thirty-nine articles over the two-year period, SPQ published two-thirds of the experiments (ten of fifteen) but none on race.

Hunt and colleagues data and my findings support the claim that social psychologists understudy race and ethnicity. Our combined results are also important for understanding why critics of other research methods have rarely trained their sights on experiments: Sociology experiments are statistically rare. Taken together, the studies raise two important questions. Why are experimental studies of race so scarce? Why should sociologists look more often to experiments to study issues in the field of race and ethnicity? In the next section, I begin answering these questions with a discussion of science and theory development.

Science and Approaches to Theory Development

Basic science has as its objectives, identifying, classifying, and explaining recurrent relationships between phenomena. Science achieves these objectives by answering an ordered series of questions as follows (Walker 2002): (1) What is the phenomenon y? (2) What phenomena (xs) are correlated with y? (3) What accounts for the relationship between y and x? Scholars answer the first type of question by introducing definitions and classification schemes. The second type of question motivates a search for relations between phenomena. Finding patterned relations can trigger the development of *empiricist* or historical explanations. Alternatively, finding patterned relations motivates researchers to ask questions of the third type. Scientists answer the third type of question by devising *theoretical* explanations. Empiricist explanations and explanations by theory reflect two different approaches to science. I describe each approach below.

Empiricist and Theory-Driven Approaches to Science

On their faces, empiricist and theory-driven approaches to science appear very similar. Each approach treats the discovery of laws or theory as the ultimate

objective of science. In each approach, experiments are also important tools in the scientist's toolkit. The approaches advocate fundamentally different strategies for creating bodies of laws.

The Empiricist Approach and Empiricist Explanations

The empiricist approach presumes that laws are found by observing the world of phenomena. Mill's five canons (1843/1967) are the logical foundations for empiricist research. The methods of (1) agreement; (2) difference; (3) agreement and difference; (4) concomitant variation; and (5) residues are often described in social science research texts. Mill's idea of an experiment is epitomized by his description of the method of difference, which, for him, is *the* experimental method:

> If an instance in which the phenomenon under investigation occurs, and an instance in which it does not occur, have every circumstance in common save one, that one occurring only in the former; the circumstance in which alone the two instances differ is the effect, or cause, or an indispensable part of the cause, of the phenomenon. (Mill 1843 /1967:452)

As the method of difference is described in most texts, it is the standard to which every other method aspires and which social researchers claim to emulate (Lieberson 1985). The "phenomenon under investigation" is an effect or dependent variable. The "circumstance in which alone the two instances differ" is the cause or independent variable. The cause occurs in the experimental group and does not appear in the control group.

Mill argued that scientific knowledge is found by applying his methods to the world of phenomena and uncovering regular patterns of cause and effect. He presumed that statements describing invariant cause-effect patterns are laws.[6] They are not. Nor can they be treated as laws. Laws are general statements that describe invariant relationships between classes of phenomena (e.g., social status and competence evaluations) and do not refer to particular events, times, or places. Here are two examples of laws:

1. "The alteration of motion is ever proportional to the motive force impressed; and is made in the direction of the right line in which that force is impressed" (Newton 1686/1966:83).
2. An actor's resistance to accepting a given exchange is equal to the ratio of (1) its interest in gaining its best payoff to; (2) its interest in avoiding its worst payoff (see Willer & Anderson [1981:122] or Willer [1999:43]).

[6]The first statement is Newton's second law of motion and, after Newton's clarifications and extensions, it has been passed down to contemporary

students as Equation 1. The idea that the quantity, force, is equal to mass times acceleration has been memorized by generations of high-school physics students.

$$f = ma \tag{1}$$

The second statement is unfamiliar to all but a few social scientists in the subfields of social psychology and exchange network studies. It is the law of resistance drawn from the Elementary Theory (Willer & Anderson 1981). The resistance law specifies the degree to which an actor, *i*, resists a proposed exchange with another, *j*. It is expressed as Equation 2:

$$R_i = \frac{P_i \max - P_i}{P_i - P_i con} \tag{2}$$

where R_i is an actor's resistance to a payoff, P_i, P_i max is *i*'s highest possible payoff, and P_i con is *i*'s payoff at confrontation when *i* fails to reach agreement with its negotiation partner, *j*. Elementary Theory (ET) and the related Network Exchange Theory (NET) are sociological theories of social structure and behavior. The resistance law is important because a principle of ET asserts that two actors exchange at the point of equiresistance (Equation 3). I shall use ideas from ET and NET below.

$$R_i = \frac{P_i \max - P_i}{P_i - P_i con} = \frac{P_j \max - P_j}{P_j - P_j con} = R_j \tag{3}$$

Neither of the two laws above is equivalent to an aggregation of observations (e.g., Observation_1, Observation_2, ..., Observation_n) because neither refers to specific events, places or times or aggregations of such events. That they are not demonstrates the failure of Mill's method.

Mill's method fails to find laws. It fails because there is no logic that permits an observer or generations of observers to infer general statements from an empirical observation or a collection of observations. Observations are described in concrete terms whereas laws and law-like sentences (Hempel 1966; Nagel 1961) employ theoretical constructs (Willer & Webster 1971).

Mill's method also fails because it is impossible to use one or a million observations to claim evidence of an absolute regularity. Observation 1,000,001 may be inconsistent with the previous million. As a result, Mill's project was modified to focus on *probable regularities* (i.e., patterns that are unlikely to occur by chance). Fisher (1935, 1956) is widely recognized for establishing the statistical foundations for inferring probable regularities from a set of concrete observations. The Mill-Fisher method is today the hallmark of academic social science and empiricist research, including empiricist experiments. Despite its failure to find laws, the Mill-Fisher approach is an excellent one for finding empirical regularities and developing empiricist explanations.

There are two types of empiricist explanation—qualitative and quantitative (Walker 2002). Qualitative empiricist explanations are understandings of a concrete event. Quantitative empiricist explanations offer understandings of collections of concrete events. Empiricist explanations are data driven; their content depends entirely on a set of putative facts.

Consider the Matewan Massacre. On May 19, 1920, a violent shootout took place in the small coal-mining town of Matewan, West Virginia (Bailey 2008). Ten people were killed and the battle triggered events that led to a coal war that took the lives of approximately fifty people over the next two years. A qualitative empiricist explanation of the massacre describes the causal connections between the gun battle (a fact) and events preceding it. Relevant events include the migration of ethnic Hungarians and blacks to the area, the reactions of old or native stock to the new residents, local political structures and actions taken by the Republican and Democrat parties, mine owners' interests and actions, and so on. These spatially and temporally constrained events and structures can only be employed to explain the Matewan Massacre.

By way of contrast, Lichter and colleagues (1997) used quantitative analysis of county-level data to explain the likelihood that families are headed by women. Their article is a textbook example of quantitative empiricist explanation. It shows that thirteen factors—including percent black—are significantly associated with the likelihood that families are headed by females. The thirteen concrete factors are said to explain variation in the dependent measure, family or household type—an aggregation of concrete events. Lichter and colleagues research, like many quantitative empiricist explanations, appears all the more impressive because it gives precise estimates of its explanatory effectiveness (measured as R^2). Their analysis explains 83.4% of variation in the dependent measure. However, like all empiricist explanations, there is less than meets the eye. Such explanations typically fail to offer insight into *why* any two phenomena are related.

The Theory-Driven Approach

As stated earlier, the theory-driven approach is concerned with developing and testing explanations by theory. The empiricist approach finds patterned relations between phenomena and presumes that the patterns are laws. In contrast, the theory-driven approach is concerned with developing theories that explain *why* two general phenomena are related. A theory is a set of interrelated, universal statements (laws or law-like statements), to which a set of rules or procedures can be applied to create new statements (Willer & Walker 2007:20). Universal statements contain constructs (i.e., general terms that do not refer to specific places or times). Consider the *abstract empirical generalization*:

For members of task groups, social status is correlated with perceptions of competence.[7]

Social status and perceptions of task-group competence are general terms. Why social status and perceptions of competence are correlated is a question that motivates theory development (Walker 2002). Theories are deductive systems that can be represented as follows:

$$\begin{aligned} &\text{E:} && {}_{A}R_{2D} \\ &\text{P1:} && {}_{A}R_{1B} \\ &\text{P2:} && {}_{B}R_{1C} \\ &\text{P3:} && {}_{C}R_{2D} \end{aligned}$$

where E is an abstract generalization requiring explanation; each P is a proposition or argument; the subscripted terms A, B, C, and D are theoretical constructs that represent phenomena (e.g., social status, competence evaluations); and R is a relation (e.g., positive association, ownership). The statements 1–3 are universal in the algebraic sense (i.e., their terms can represent any phenomena or relations). P1 is interpreted as "A stands in a particular relation, R_1, to B," and similarly for P2 and P3. The representation includes R_1 and R_2 because theories often include statements about different types of relations.[8]

The theory of Status Characteristics and Expectation States (SCT) offers an explanation for the generalization above (Berger et al. 1977). The theory consists of interconnected propositions or arguments that explain the connection between social status and perceptions of competence (e.g., "A" and "D" in the schema above).[9] As a general theory, SCT can explain a host of phenomena including the correlation between race/ethnic status and perceptions of competence, sexual orientation and perceptions of competence, and so on.

Theories are data generating rather than data dependent because they must be tested against empirical reality. To do so, a researcher first uses the theory to build a model. Next, she or he identifies measures of: (1) a theory's concepts; (2) conditions that realize the theory's scope of application; and (3) its initial conditions to create a research design that generates data. A research design is a replica of its theoretical model (Freese & Sell 1980). Data that fail to fit hypotheses drawn from theory put the theory in danger of falsification (Popper 1934/1958). Theories that fail to find empirical support from carefully designed studies must be revised or discarded. Data that fit hypotheses deduced from a theory support the theory's conditional truth.[10] Theories that find empirical support can be refined and improved to explain a broader range of phenomena or to explain phenomena more precisely. Figure 7.1 describes theory-driven research and the first phases of a strategy for developing cumulative theoretical knowledge.

Important differences in the empiricist and theory-driven approaches to science are reflected by research carried out under those approaches. I discuss research and experiments as a research method in the next section.

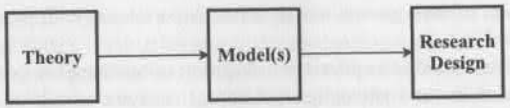

Fig. 7.1 Relations between theory, model, and research design

Research: Objectives and Techniques

Two types of research follow from the empiricist and theory-driven approaches to science. Empiricist research uncovers phenomena and triggers the development of empiricist explanations. Theory-driven research tests theories. Explanations by theory can exist independently of data, but *theory testing* requires data and that requirement motivates research design and data collection (see Figure 7.1). There is an important implication of the research process: Any evaluation of research techniques must include evaluations of the approach taken as well as the soundness of designs.

Empiricist research falls into two broad categories. The first is exploratory research that involves a search for patterned behavior. It is not guided by hypotheses in any meaningful sense of that term. The second type of empiricist research is the de facto standard for sociological research. It tests hypotheses drawn from hunches and extrapolations from previously published research. As stated above, experiments are the ideal method for research conducted in the Mill-Fisher approach.

Theory-driven research is very different than empiricist research. Its objective is to test hypotheses drawn from research designs that, in turn, are drawn from theoretic models. Theory-driven research is the key to scientific advance, and experiments are the best method for testing theory. I turn now to experiments and experimental research.

What Is an Experiment?[11]

An experiment is "an inquiry for which the investigator controls the phenomena of interest and sets the conditions under which they are observed and measured" (Willer & Walker 2007:2). Experiments are rarely used in sociology. There are several reasons why this is so. First, sociologists-in-training rarely learn the purposes of experiments or how to design and conduct them. (See Webster & Sell [2007] for an excellent introduction to experiments in the social sciences.) Second, poor training promotes and sustains popular misconceptions about experiments. Many sociologists believe that laboratory experiments are only useful for studying microsocial phenomena. Third, many sociologists believe or have been taught that the *artificiality* of experiments makes generalizing experimental findings difficult if not impossible.

Fortunately, all of these barriers to increasing the numbers of experiments that study race and ethnicity are surmountable.

Training

Courses in research methods and statistics typically offer sociologists-in-training limited exposure to experimental methods. Students learn that

> The most conventional type of experiment, in the natural as well as the social sciences, involves three major pairs of components: (1) independent and dependent variables, (2) pretesting and posttesting, and (3) experiment and control groups. (Babbie 1998:233)

> *True experiments* must have at least three things: Two comparison groups (in the simplest case, an experimental and a control group), variation in the independent variable before assessment of change in the dependent variable, [and] random assignment to the two (or more) comparison groups. (Schutt 2006:201; emphasis in the original)

These descriptions and those in other textbooks focus on *empiricist* experiments, the standard model for sociology experiments. The Mill-Fisher empiricist experiment can be summarized by four maxims that mirror the definitions of empiricist experiments given above. (See Willer and Walker [2007] for a more detailed discussion and illustration of each maxim.)

1. Create at least two study conditions that are initially as identical as possible.
2. Introduce a single difference (a presumed cause) between the two conditions and observe the result.
3. Restrict all inferences about the result to the effect(s) of that single difference.
4. Infer relative regularities only if it is unlikely that they are due to chance.

The empiricist experiment as refined by Fisher is a very effective tool for uncovering and verifying probable regularities if research conforms to the maxims above.

Beginning students are rarely exposed to theory-driven experiments—a form known from antiquity and institutionalized in the physical sciences since Galileo (Willer & Walker 2007:12 ff.). Theory-driven experiments test theory. *Theory designs methods of test* that generate evidence for or against a theory's knowledge claims. Theories that garner substantial empirical support can be applied for prediction or explanation. The design of theory-driven experiments is given by the following maxims:

1. Derive one or more models from the theory to be tested.
2. Use the theory to generate predictions by linking initial conditions to end conditions.
3. Build experiment replicas, set initial conditions, and observe the end conditions.
4. Compare results to predictions and decide whether the theory is supported.
5. Make inferences from theory with greatest confidence to instances most theoretically similar to experiments supporting the theory. Predictions are not formally limited by that similarity. (Again, see Willer and Walker [2007] for detailed discussion of these maxims.)

Experiments are the best methods for testing hypotheses and theory because they give researchers the greatest control over test conditions. Experimental control serves different purposes in empiricist and theory-driven experiments. Empiricist experiments control test conditions in order to create experimental and control conditions that are as similar as possible. The Mill-Fisher method requires their essential similarity to infer that putative causes are probable causes of effects.

Theory-driven experiments control test conditions to create conditions that satisfy conditions described by theory. Put another way, test conditions are controlled by (i.e., established by) theory so that results can be credibly evaluated as supporting or disconfirming the theory. Because theory designs the experiments, theory-driven experiments will have designs that reproduce—in concrete form—the general relationships found in the theories they test. It follows that the designs of theory-driven experiments will vary as widely as the theories that are tested.

Social Science Experiments with Applications to Race and Ethnic Studies

Varieties of Experimental Designs

Scholars can use research to uncover patterned relations or to test theory. Empiricist research can be used to study any phenomenon that piques the spirit of human inquiry. Theory-driven research can be used to study any phenomena that are explained by theory. Researcher control is the criterion that distinguishes experiments from other research techniques. Laboratory experiments offer researchers the greatest control of study conditions but there are other experimental designs. In what follows, I give examples of natural, field, survey, and laboratory experiments. Some examples are hybrids of pure types but the sample demonstrates the wide range of issues that can be investigated experimentally.

Natural Experiments

There are instances in which the world of phenomena offers researchers conditions that imitate experimental control. Natural experiments can produce interesting discoveries and, in some cases, offer unusual tests of theory. The draft lottery implemented in 1969 spawned several natural experiments. The Selective Service System randomly assigned a number to each day of the calendar year to establish the draft order for eligible men. Draft-eligible males born on June 8 "won" the first lottery held in December 1969. The date was assigned the number 366, whereas those born on September 14 got number 1. Men with low lottery numbers were drafted first, so the lottery neatly partitioned more than 5 million men into two groups, those certain to be drafted and those almost certain to remain undrafted. Alternatively, draft numbers can be treated as continuous measures of exposure to the draft.

A decade after the Vietnam War ended, Hearst et al. (1986) reported that men with low draft numbers were more likely than those with higher numbers to commit suicide and to die in traffic accidents. The findings were interesting, provocative and were used to justify providing increased medical and social support services for Vietnam-era veterans. A quarter century later, Conley and Heerwig (2009) reported findings from a draft lottery experiment that included race as an independent variable. Their study of more than 372,000 deceased members of the 1950–1952 birth cohorts failed to find a draft lottery effect. Their finding is robust and consistent across race and ethnic subgroups (white, non-Hispanic black, Hispanic, and non-Hispanic others). Conley and Heerwig conclude that special assistance for health effects of draft status may not be needed over the long term. Hearst and colleagues cannot explain why there appear to be short-term draft-lottery effects on health, nor can they or Conley and Heerwig explain why the effects disappeared two decades later. Answering those questions requires theoretical analysis.

Field Experiments

Field experiments are conducted in natural settings and exercise greater control over research situations than naturalistic observations (e.g., ethnography or participant observation). Audit studies that use testing teams or telephone audits (Massey & Lundy 2001) to detect housing discrimination are field experiments. LaPiere's (1934) study of attitudes and actions is a classic field experiment. LaPiere traveled around the country with a Chinese couple visiting 251 hotels, restaurants, and campgrounds. The group was refused service at one establishment. Six months after the last visit, LaPiere mailed the proprietors and asked if they would serve Chinese, Negroes, or other ethnic patrons. Of 128 responses, only one proprietor responded with an unqualified "yes" to the query about serving Chinese patrons. The findings show extremely poor fit between attitudes (prejudice) and actions (discrimination). LaPiere could not

explain his findings but speculated that subtle and uncontrolled differences in demeanor (e.g., the couple spoke impeccable English and smiled a lot), not ethnicity, may have determined proprietors' positive reception of the traveling party. Ayres and Siegelman (1995) offer similar post hoc speculation concerning findings of their field experiment in which white males got better offers for new car purchases than black males or black and white females.

Survey Experiments

Survey experiments control the content of survey items and the order in which they are presented. Respondents often differ on many characteristics and are typically drawn from a variety of social and personal situations. Most survey experiments are empiricist studies that apply the Mill-Fisher paradigm in an efficient and effective way. However, some survey experiments are used to test theory.

Krysan's (1998) research on white racial attitudes is an example of a survey experiment. Krysan began with an apparent paradox. Surveys show that white support for principles of race and ethnic equality has risen over time. Yet, the same body of research finds declining white support of policies that are proposed to ameliorate race and ethnic inequalities. Krysan speculated that normative pressure (independent variable) influences public expressions of racial attitudes (dependent variable) and designed a survey experiment to evaluate the idea. The survey included traditional items used to measure race prejudice like support of a black presidential candidate, support of intermarriage of Whites with members of various race or ethnic groups, tolerance of black neighbors, attributions of black disadvantage to social or personal (i.e., innate) conditions, etc. Krysan asked white respondents to answer questions under three experimental treatments as follows:

1. Public: Respondents answered questions in face-to-face interviews (high pressure).
2. Semi-private: Respondents answered some questions in face-to-face interviews but were interrupted periodically to complete some paper-and-pencil items (moderate pressure).
3. Private: Respondents answered survey questions mailed to their homes (low pressure).

Krysan reported partial support for her first hypothesis that white respondents would express less liberal attitudes as the privacy of their responses increased. The positive finding did not hold for some attitude items. Her second hypothesis that privacy effects would be greater for items about principles of racial equality than items about racial policies was disconfirmed. She also found mixed support for a third hypothesis that privacy effects were more

pronounced for the highly educated. College graduates who responded privately expressed more negative attitudes than other combinations of education and mode of administration but, again, only for some items.

Krysan's research addressed issues of theoretical and practical importance but she was unable to find strong support for the three regularities she expected to find. As a next step, the findings of empiricist experiments like this one suggest mounting additional studies to answer conclusively whether there is a regular relationship between privacy and expressions of race tolerance.

Laboratory Experiments

Laboratory experiments are conducted under highly controlled conditions. Researchers exercise maximal control of important variables for empiricist and theory-driven experiments. Here, I describe an empiricist laboratory experiment and a series of theory-driven experiments that demonstrate the range of phenomena that can be investigated with theory.

The Obama Effect

A series of recent papers focuses on what some label the "Obama effect." The idea is that the success of Barack Obama's presidential campaign has had or can have positive effects on the aspirations and behaviors of black Americans. Several experiments draw on the concept of *stereotype threat*, a reaction that individuals experience when evaluations of their performances have the potential to verify negative stereotypes of their ingroup (Steele & Aronson 1995). Stereotype threat is assumed to degrade the performances of those who experience it.

Aronson et al. (2009) used a laboratory experiment to study the Obama effect. In July 2008, they administered portions of the Medical College Admission Test (MCAT) to aspiring medical students enrolled in summer programs at three universities. The experiments were conducted before the political conventions had officially designated Barack Obama and John McCain as presidential candidates. All participants received an instruction that described the test as an important measure of ability that predicted success in gaining admission to either medical or graduate school. That instruction was designed to activate stereotype threat.

Students were randomly assigned to one of four treatments: (1) Obama treatment; (2) McCain treatment; (3) American politician treatment; and (4) control. Subjects in each of the four groups were given a test booklet that contained a cover sheet, test instructions, and the MCAT items. Test booklets for treatments 1–3 included a short survey of political issues designed to get subjects to think about and offer positive comments about Obama, McCain, or "an American politician." Participants in the Obama and McCain conditions saw three small color photographs of the candidate at the top of the sheet

(e.g., smiling on a *Time* magazine cover). A series of politically ambiguous quotations taken from speeches, news conferences, etc., was printed under the photos. The quotations were attributed to Obama, McCain, or to "an American politician." The quotes were identical with one exception. The phrase "my friends," uniquely identified with McCain during the campaign, was added when the quotes were attributed to McCain.

Aronson et al. (2009) tested the hypothesis that black students who are primed to think positively about Obama before taking a difficult test will have higher scores than Blacks who are not primed. Their hypothesis was not supported. Blacks in the Obama treatment did not have higher test scores than Blacks in the other treatments. Further analysis showed that test performances were not affected by the degree of positive sentiment expressed toward Obama, engagement with the presidential campaign, mentioning Obama's race, mentioning Obama's intelligence, or the subjects' SAT scores. In fact, students who showed greater engagement with the election (measured by how closely the respondent followed events on television during the previous three months) had lower test scores.

Aronson and colleagues findings are interesting but answer no important questions. There is no evidence of an Obama effect and their findings are inconsistent with those of Marx et al. (2009), who asked 472 adult Americans to answer items from the Graduate Record Exam. Marx and colleagues predicted that Obama's candidacy could reduce black-white differences generated by stereotype threat. They report smaller black-white differences in test scores immediately after "stereotype-defying" actions during the Obama campaign. Differences in the studies' findings are intriguing, but strict comparisons are not possible because the experiments differ in many ways. Marx and colleagues study combined elements of a natural experiment with typical laboratory procedures. It was a web-based study in which respondents answered items under two conditions (relatively active or relatively inactive periods of the Obama campaign). Aronson and colleagues research was conducted in labs at three universities with students contemplating medical careers. Is there an Obama effect? These studies leave the basic question unanswered.

But Is It a Real World? Theory, Micro Experiments, and Macro Application

Two related criticisms are often directed toward social science experiments. First, some critics argue that research labs are contrived social situations and that findings from experiments cannot be generalized beyond the lab. Second, given constraints on space, laboratory groups are small. As a result, even if they could be applied outside the laboratory, findings from such studies would not be useful for understanding meso- and macro-level phenomena. Each

criticism is important but neither is an impediment to using experiments as an integral part of the scientific process, nor are they barriers to scientific progress. In fact, the features of experiments that appear most problematic make experiments the most suitable method for theory-driven science.

Laboratory experiments that test theory are contrived situations but they are also real situations. Researchers can exercise the greatest level of control in the lab. They use theory to build situations that are stripped of the baggage found in natural settings—baggage that proves detrimental to tests of theoretical understandings of social life. The claim that their artificiality makes it impossible to generalize laboratory findings misunderstands the role of data in the scientific process. As stated earlier, there is no logic that permits an investigator to generalize one or a million research findings to find laws as Mill presumed. Nor can findings (e.g., sample data) be generalized to a population. That reality is devastating for scholars who take an empiricist approach to science, but generalization, as espoused by empiricists, is irrelevant to theory-driven research.

Theory-driven research does not try to find laws. It tests laws and the theories in which they are embedded. Moreover, theory-driven research is not concerned with applying patterns observed in the lab to situations beyond the research setting. For theory-driven research, generalization is a process through which *theory* is applied more generally, first within a specified scope and, as scope restrictions are relaxed, over an ever-expanding scope.

Physical constraints limit the size and range of social systems created in research labs. Most laboratory situations do not qualify as microcosms of any known universe. In that sense, uninformed critics have every reason to be concerned about using experiments to study meso- and macro-level phenomena. Armies cannot be brought into research labs (Zelditch 1969). Some critics misunderstand that elemental truth to mean that researchers cannot use experiments to study organizational life or life in other large-scale social institutions. Nothing could be further from the truth. Researchers can use laboratory experiments to gain vital information about processes that play out among military units or between nation states (Zelditch 1969). I use experiments that test predictions of Elementary Theory (ET) to illustrate.

As mentioned previously, ET (Willer & Anderson 1981) is a theory of network structures and social behavior that describes and explains processes of exchange, coercion, and conflict. The experiments I describe below use the coercive structure reproduced as Figure 7.2. The *D*s possess resources that are of value to them and to *C*. *C* is an actor who controls negative sanctions that can be used against actors at D_1–D_4. As is true of the "real world," *C*'s decision to use sanctions is costly.

Experimenters use student volunteers to fill the positions in Figure 7.2. Subjects are given chips at the start of each experiment trial. Position *C* is given 4 red chips that it can send to any or all *D*s. Each *D* is given 10 white

Application of ET to the study of slavery in Brazil and the United States is more than an exercise. ET predicts that harsh sanctioning of slaves (*D*s) will increase their marginal productivity as they compete to avoid punishment. On the other hand, confrontation is costly to slaves and slave holders (*C*s). Slaves must bear the psychic costs associated with living under threat of bodily harm and the physical costs of severe injury or death. Slave holders lose the value of dead or incapacitated slaves and must incur the costs of replacement. The costs of confrontation for slaves (p_d con) are assumed to be relatively stable across slaving societies, but a plentiful supply of slaves lowers the cost of replacing them and reduces the cost of confrontation (p_c con) for slave holders. Consequently, p_c con is lower in countries that had plentiful supplies of slaves (e.g., Brazil between 1808 and 1888) than in countries like the United States where the supply was less plentiful. Application of ET (Equations 2 and 3) leads to the twin predictions that slaves are treated more harshly and their labor exploited to produce greater profits in countries where supplies of slaves are bountiful. ET's predictions are consistent with historical observations and are based on well-supported theory rather than informed speculations like those of Marx and Weber.

Concluding Remarks and an Appeal

Theory is underdeveloped and researchers rarely conduct experiments in race and ethnic studies. This chapter advocates a shift from empiricist to theory-driven research as a corrective for both problems. Theory development is the ultimate goal of basic science, and theory-driven research is more productive of theory. Theory designs methods of test and experiments are the best method for testing, given researchers' ability to exercise control of the research setting. Additionally, researchers properly trained in experimental techniques can more easily apply tested theories to situations outside the lab. Any discipline that amasses a larger storehouse of theory than another also has an advantage in using theory—rather than ill-informed speculations—to design policies. Better theoretical understandings increase the odds that policies have positive results (i.e., produce intended outcomes). There are other salutary effects of using a theory-driven approach that develops more theory and makes more extensive use of experiments.

Critiques of race and ethnic studies find myriad problems. Many studies reflect various forms of race and cultural bias. Bias taints the research process by leading researchers to misperceive phenomena, misinterpret observations, and offer explanations for phenomena that either do not exist or have been described inaccurately by biased observers. Some bias follows from the underrepresentation and underutilization of researchers and research subjects who are members of race and ethnic minority groups. All too often, researchers

have made observations of dominant race and ethnic groups and *assumed* that the same patterns ought to be reflected among members of subordinate groups. The failure of some groups to exhibit expected patterns has often been treated as an aberration to be accounted for by empiricist explanations. Finally, cultural and race prejudices contribute to malfeasance and maltreatment of minority participants in research and the communities from which they are drawn.

There are many valid criticisms of specific race and ethnic studies, and I applaud those who point them out. However, I raise several general concerns about critical commentary. First, critics all too often fail to identify correctly the sources of the discipline's troubles. Second, few of the criticisms are constructive. Third, our discipline and its practitioners have not addressed clearly and specifically the difficulties Du Bois associated with finding and applying scientific laws. I address these objections in turn.

Many of the most vocal critics of research in our field attribute problems to the logic and origins of modern science and research methods. However, since Du Bois (1898), it has been clear that the actual culprits are social scientists who misunderstand or misuse research techniques and scientific method. Neither the logic of modern science nor the research methods scientists use can be held accountable for underrepresentation of minority scholars, bias in research, or malfeasance of researchers and research institutions. However, the approach a scientist takes can exacerbate or ameliorate problems identified by critics.

My second objection is to critics who describe problems but don't offer plausible solutions. If solutions are offered, they are often based on untested hypotheses or are so vague that they are useless as guides to corrective action. Consider criticisms of the Moynihan (1965) report. Moynihan reported a negative statistical association between the number of adult heads of families and family poverty. Critics claimed that what Moynihan saw as weaknesses were really strengths when viewed from a minority rather than majority perspective (Billingsley 1968; Hill 1972). However, the critics failed to describe procedures that would produce a set of contrary findings (e.g., a zero or positive correlation).

My final and most strenuous objection concerns the failure of sociologists—not just scholars of race—to fully specify and address adequately Du Bois's (1940) concerns about "difficulties of applying scientific law and discovering cause and effect in the social world." In contrast, I argue that lagging theory development is due in part to a reliance on an empiricist approach to science. There is an additional disadvantage of the empiricist approach. The current period is one in which *public sociology* is appealing to a growing proportion of sociologists (Burawoy 2005). Public sociology incorporates a number of ideas, but one of them reflects Du Bois's interest in applying sociological knowledge to public issues. I agree with Weber (1918/1958) that

speaking out—and taking action—on public issues is our "damned duty." At the same time, I argue that mixing social activism and science creates bad results.

Taking active positions on social issues requires activating and acting on values. Activists value some social arrangements more than others (e.g., integrated vs. segregated schools). Activists are more likely to value observing the correlates of outcomes (empiricism) rather than explaining relations between outcomes and their correlates (theory). Injecting values into scientific work ensures that sociological research is affected by bias. A half-century has passed since Rossi (1960) asserted that without overarching theory, investigators typically find exactly what they are looking for rather than what exists. Making public whose side we are on (Becker 1966) does nothing to mitigate the negative effects of activist bias.

There is a way out of the conundrum, but it requires embracing science—the enemy many critics have named. Sociologists of race and ethnicity can lead the way by reforming not science but scientists. To do so, scientists must embrace theory—the method of science. Not method as investigative technique but method as the logic that justifies accepting or rejecting theory. We can make headway in the fight to resolve problems raised by critics if we recognize that doing science is a public process. As scientists we comprise a public community, and external funding exposes our actions to a much larger public. Only public scrutiny can prevent abuses like some of those from the past. The Tuskegee Experiment (Jones 1981) was immoral and unethical, and ethical safeguards introduced after it was brought to light make future Tuskegees less likely. But we must be vigilant. Today, many people are repulsed by the idea of harvesting embryonic stem cells from aborted fetuses just as an earlier generation was repulsed by revelations of the Tuskegee study. Women from some groups are disproportionately more likely to have abortions—a potential source of stem cell material—than are members of the majority. Scholars of race can play an important role in regulating the actions of scientists.

On the main criticisms of race and ethnic research, theory and data collection are objects of scrutiny by the scientific and lay communities. Theory can be tested by anyone with the requisite knowledge, skills, and resources to do so. Interested and independent analysts can check data that purport to support theory. It is still possible to claim that American Blacks do less well academically than Whites because they are genetically inclined to poorer performances. And some researchers continue to make such claims. But theories that purport to explain how genetic characteristics produce poorer performances must be tested. And tests can be conducted by qualified members of the groups that have been described as genetically inferior.

The distinction between empiricist and theory-driven experiments raises a caution. Experimenter control is the hallmark of both types of research but the nature and purpose of control is different. Empiricist experiments seek

answers to "which *x*s are related to *y*" questions. Aronson et al. (2009) asked whether seeing positive images and making positive statements about Barack Obama (x) overcame the negative effects of stereotype threat (y). They used experimental control to isolate key independent and dependent variables. Using standard Mill-Fisher logic, they examined the question and found no support for their hypothesis.

Theory-driven studies rely on theory to design experimental controls. Theory-driven experiments are designs that establish conditions that satisfy a theory's scope and initial conditions and use reliable and valid measures of a theory's concepts. For theory-driven experiments, experimental control means *controlled by theory*. Willer (1987) created coercive structures in his laboratory because ET required them. The theory predicts payoffs for *C* and *D* in lab studies and for master-slave relations in slave-holding societies. The theory is supported by observations from the lab and from two slave-holding societies.

Empiricist and theory-driven research can be complementary. Scholars can repeat Aronson et al.'s (2009) or Willer's (1987) experiments. Perhaps, members of Aronson and colleagues research team supported candidates other than Obama and introduced bias into the study. The public character of the research process permits others to replicate Aronson and colleagues empiricist experiment to find evidence of an Obama effect obscured by bias. After uncovering an effect, researchers would be left to contemplate the next step. Does existing theory explain the Obama effect? If so, patterns found by empiricist research can fuel theory-driven research to replicate and extend the findings. If no theory explains the now-documented phenomenon, science dictates that theorists take up the challenge to build new theory and test it as part of a theory-driven program. But all of this is part of a public process—a process that abhors and corrects bias, misrepresentation, and bad ideas.

The theory-driven approach values theoretical understandings. Theories that find support in the lab can be applied to a variety of micro-, meso-, and macro-level phenomena that satisfy the theory's conditions. Given the proven value of theory-driven lab experiments, scholars of race ought to be more inclined to use them in their research projects. Moreover, findings from such research can be used to inform public policy as "others will." One can only hope that Du Bois would approve.

NOTES

1. Following Rudner (1966:5), I distinguish techniques for collecting and analyzing data (method$_1$) from the logic of justification—a discipline's rationale for accepting or rejecting theories (method$_2$ or methodology). Here and elsewhere, I use the terms "methods" and "techniques" interchangeably. I trust that I am able to communicate clearly whether the terms are intended to refer to method$_1$ or method$_2$.

2. The term "empiricist" is often used as a pejorative. I use the word in its descriptive sense (i.e., that knowledge [theory] can be derived from observations).
3. Ernest Manheim was a remarkable man and brilliant teacher. Fearful that the Nazis would not approve it, Manheim withdrew his habilitation thesis (an academic qualification in Germany that comes after the Ph.D.) in 1933 shortly after the faculty at Leipzig had approved it. He moved to London and worked as an assistant to his cousin, Karl Mannheim, while earning a doctorate in anthropology. Arriving at Kansas City University (later UMKC) in 1938, his research on race relations quickly reshaped the thinking of civic leaders in Kansas City. Later, Manheim volunteered to testify for Brown when *Brown v. Topeka* was heard in federal district court.
4. The volume of criticisms has expanded almost as rapidly as the field. See papers compiled by Ladner (1973), Stanfield and Dennis (1993), and Zuberi and Bonilla-Silva (2008) for a thorough introduction to general criticisms of research on race and ethnicity.
5. SF and AJS split volumes across calendar years. I reviewed articles published in the appropriate volumes for 2006–2008. AJS, SF, and SPQ also published special issues that focused on race during the period of observation. Consequently, the numbers of papers devoted to race may be inflated in those journals. Finally, as expected, all articles in ERS focus on race.
6. An empiricist approach "emphasizes the importance of observation and of creating knowledge by amassing observations and generalizing from these observations" (Cohen 1989:16; and see Popper 1962:21ff.). Mill's logic is empiricist.
7. Examples of empirical generalizations are: (1) Mexican American students earn lower grades than non-Hispanic white Americans; or (2) students give higher evaluations to attractive than to unattractive teachers (Freng & Webber 2009; Hamermesh & Parker 2005). The claim that perceptions of task competence vary with social status is an abstract empirical generalization. The statement describes a relationship between two general concepts (or constructs), social status and perceived task competence, where Mexican American, non-Hispanic white, and attractiveness are examples of status groups, and grades and student evaluations of teaching are perceptions of competence.
8. I simplify the description of theory for illustrative purposes. Willer and Walker (2007:17–30) give an introduction to the structure of theories and Cohen (1989) offers an extended introduction suitable for advanced undergraduates, graduate students, and others who seek a nontechnical introduction to theory construction.
9. SCT consists of five propositions and a number of scope restrictions. I do not present the theory here, but see Berger et al. (1977) for a complete discussion of the theory, including arguments and scope restrictions. See Walker (1999) for an application of SCT to modern affirmative action and interracial relations.
10. Theories are never judged true in an absolute sense. Tests of theories are not tests of reality but tests of sentences that claim to describe reality. Even well-corroborated theories are in danger of falsification if they encounter contrary evidence in future tests. For this reason, it is important that theorists specify the scope limitations of their formulations (Walker & Cohen 1985) and embed tests of theory in a larger program of theory-driven research (Cohen 1997).
11. Remarks in this section draw extensively on my previous work (Walker 2002) and work with David Willer (Walker & Willer 2007; Willer & Walker 2007).

REFERENCES

Allport, G. (1954). *The nature of prejudice*. New York: Addison-Wesley.

Aronson, J., Jannone, S., McGlone, M., & Johnson-Campbell, T. (2009). The Obama effect: An experimental test. *Journal of Experimental Social Psychology, 45*, 957–960.

Ayres, I., & Siegelman, P. (1995). Race and gender discrimination in bargaining for a new car. *American Economic Review, 85*, 304–321.

Babbie, E. (1998). *The practice of social research*. Belmont, CA: Wadsworth.

Bailey, R. J. (2008). *Matewan before the massacre*. Morgantown: West Virginia University Press.

Becker, H. S. (1966). Whose side are we on? *Social Problems, 14*, 239–247.

Berger, J., Fisek, M. H., Norman, R. Z., & Zelditch, M., Jr. (1977). *Status characteristics and social interaction*. New York: Elsevier.

Billingsley, A. (1968). *Black families in white America*. Englewood Cliffs, NJ: Prentice-Hall.

Burawoy, M. (2005). For public sociology. *American Sociological Review, 70*, 4–28.

Clark, W. A. V. (2002). Ethnic preferences and ethnic perceptions in multi-ethnic settings. *Urban Geography, 23*, 237–256.

Cohen, B. P. (1989). *Developing sociological knowledge*, 2nd ed. Chicago: Nelson-Hall.

Cohen, B. P. (1997). Beyond experimental inference: A decent burial for J. S. Mill and R. A. Fisher. In J. Szmatka, J. Skvoretz, and J. Berger (Eds.), *Status, network and structure: Theory development in group processes* (pp. 71–86). Stanford, CA: Stanford University Press.

Cole, S. (Ed.) et al. (1994). Special issue: What's wrong with sociology. *Sociological Forum, 9*, 129–320. New York: Plenum Press.

Conley, D., & Heerwig, J. A. (2009). The long-term effects of military conscription on mortality: Estimates from the Vietnam-era draft lottery. Cambridge, MA: National Bureau of Economic Research, Working Paper 15105.

Degler, C. N. (1971). *Neither black nor white*. New York: Macmillan.

Du Bois, W. E. B. (1898). The study of the Negro problems. *Annals of the American Academy of Political and Social Science, 11*, 1–23.

Du Bois, W. E. B. (1940). *Dusk of dawn*. New York: Harcourt Brace.

Fisher, R. A. (1935). *The design of experiments*. London: Oliver and Boyd.

Fisher, R. A. (1956). *Statistical methods and scientific inference*. Edinburgh: Oliver and Boyd.

Freese, L., & Sell, J. (1980). Constructing axiomatic theories in sociology. In L. Freese (Ed.), *Theoretical methods in sociology: Seven essays* (pp. 263–368). Pittsburgh: University of Pittsburgh Press.

Freng, S., & Webber, D. (2009). Turning up the heat on online teaching evaluations: Does "hotness" matter? *Teaching of Psychology, 36*, 189–193.

Goar, C. (2008). Experiments in black and white: Power and privilege in experimental methodology. In T. Zuberi & E. Bonilla-Silva (Eds.), *White logic, white methods* (pp. 153–162). Lanham, MD: Rowman & Littlefield.

Graves, J. L., Jr. (2004). *The race myth*. New York: Dutton.

Green, D. S., & Driver, E. D. (1976). W. E. B. Du Bois: A case in the sociology of sociological negation. *Phylon, 37*, 308–333.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74*, 1464–1480.

Hamermesh, D. S., & Parker, A. (2005). Beauty in the classroom: Instructors' pulchritude and putative pedagogical productivity. *Economics of Education Review, 24*, 369–376.

Hearst, N., Newman, T. B., & Hulley, S. B. (1986). Delayed effects of the military draft on mortality: A randomized natural experiment. *New England Journal of Medicine, 314*, 620–624.

Hempel, C. G. (1966). *Philosophy of social science*. Englewood Cliffs, NJ: Prentice-Hall.

Hempel, C. G., & Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science, 15*, 135–175.

Hill, R. B. (1972). *The strengths of black families*. New York: Emerson Hall.

Hunt, M. O., Jackson, P. B., Powell, B., & Steelman, L. C. (2000). Color-blind: The treatment of race and ethnicity in social psychology. *Social Psychology Quarterly, 63*, 352–364.

Jones, J. H. (1981). *Bad blood: The Tuskegee syphilis experiment*. New York: Free Press.

Kaplan, A. (1964). *The conduct of inquiry*. San Francisco: Chandler Publishing.

Krysan, M. (1998). Privacy and the expression of white racial attitudes. *Public Opinion Quarterly, 62*, 506–544.

Ladner, J. A. (1973). *The death of white sociology*. New York: Random House.

LaPiere, R. T. (1934). Attitudes vs. actions. *Social Forces, 13*, 230–237.

Lichter, D. T., McLaughlin, D. K., & Ribar, D. C. (1997). Welfare and the rise in female-headed families. *American Journal of Sociology, 103*, 112–143.

Lieberson, S. (1985). *Making it count*. Berkeley: University of California Press.

Mann, C. R. (1993). *Unequal justice: A question of color*. Bloomington: Indiana University Press.

Marx, D. M., Ko, S. J., & Friedman, R. A. (2009). The "Obama Effect": How a salient role model reduces race-based performance differences. *Journal of Experimental Social Psychology, 45*, 953–956.

Marx, K. (1867/1967). *Capital*. New York: International Publishers.

Massey, D. S., & Lundy, G. (2001). Use of black English and racial discrimination in urban housing markets: New methods and findings. *Urban Affairs Review, 36*, 452–469.

Mill, J. S. (1843/1967). *A system of logic*. London: Longmans, Green and Co.

Montagu, A. (1942). *Man's most dangerous myth: The fallacy of race*. New York: Columbia University Press.

Moynihan, D. P. (1965). *The Negro family: The case for national action*. Washington, DC: Government Printing Office.

Nagel, E. (1961). *The structure of science*. New York: Harcourt, Brace & World.

Newton, I. (1686/1966). *Principia mathematica*. Trans. A Motte. Berkeley: University of California Press.

Ogbu, J. U. (2003). *Black American students in an affluent suburb*. Mahwah, NJ: Lawrence Erlbaum.

Patterson, O. (1982). *Slavery and social death: A comparative study*. Cambridge, MA: Harvard University Press.

Popper, K. R. (1934/1958). *The logic of scientific discovery*. New York: Basic Books.

Popper, K. R. (1962). *Conjectures and refutations*. New York: Basic Books.

Rossi, P. H. (1960). Power and community structure. *Midwest Journal of Political Science, 4*, 390–401.

Rudner, R. S. (1966). *Philosophy of social science*. Englewood Cliffs, NJ: Prentice-Hall.

Schutt, R. K. (2006). *Investigating the social world: The process and practice of research*. 5th ed. Thousand Oaks, CA: Sage.

Stanfield, J. H., II, & Dennis, R. M. (1993). *Race and ethnicity in research methods*. Newbury Park, CA: Sage.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African-Americans. *Journal of Personality and Social Psychology, 69*, 797–811.

Walker, H. A. (1999). Two faces of diversity: Recreating the stranger next door? In P. Moen, D. Dempster-McClain, & H. A. Walker (Eds.), *A nation divided: Diversity, inequality and community in American society* (pp. 52–69). Ithaca, NY: Cornell University Press.

Walker, H. A. (2002). Three faces of explanation: A strategy for building cumulative knowledge. In J. Szmatka, M. Lovaglia, & K. Wysienska (Eds.), *The growth of social knowledge* (pp. 15–31). Westport, CT: Praeger.

Walker, H. A., & Cohen, B. P. (1985). Scope statements: Imperatives for evaluating theory. *American Sociological Review, 50*, 288–301.

Walker, H. A., & Willer, D. (2007). Experiments and the science of sociology. In M. Webster, Jr. & J. Sell (Eds.), *Laboratory experiments in the social sciences* (pp. 25–55). Burlington, MA: Academic Press.

Weber, M. (1896/1963). The social causes of the decay of ancient civilizations. In R. Kahl (Ed.), *Studies in explanation* (R. Frank, Trans.) (pp. 339–356). Englewood Cliffs, NJ: Prentice-Hall.