INTRODUCTION

CHAPTER 1

Experimentation in Political Science

James N. Druckman, Donald P. Green, James H. Kuklinski, and Arthur Lupia

In his 1909 American Political Science Association presidential address, A. Lawrence Lowell (1910) advised the fledgling discipline against following the model of the natural sciences: "We are limited by the impossibility of experiment. Politics is an observational, not an experimental science . . . " (7). The lopsided ratio of observational to experimental studies in political science, over the one hundred years since Lowell's statement, arguably affirms his assessment. The next hundred years are likely to be different. The number and influence of experimental studies are growing rapidly as political scientists discover ways of using experimental techniques to illuminate political phenomena.

The growing interest in experimentation reflects the increasing value that the discipline places on causal inference and empirically guided theoretical refinement. Experiments facilitate causal inference through the transparency and content of their procedures, most notably the random assignment of observations (a.k.a. subjects or experimental participants) to treatment and control groups. Experiments also guide theoretical development by providing a means for pinpointing the effects of institutional rules, preference configurations, and other contextual factors that might be difficult to assess using other forms of inference. Most of all, experiments guide theory by providing stubborn facts – that is, reliable information about cause and effect that inspires and constrains theory.

Experiments bring new opportunities for inference along with new methodological challenges. The goal of the *Cambridge Handbook of Experimental Political Science* is to help scholars more effectively pursue experimental opportunities while better understanding the challenges. To accomplish this goal, the *Handbook* offers a review of basic definitions and concepts, compares experiments with other forms of inference in political science, reviews the contributions of experimental research, and presents important methodological issues. It is our hope that discussing these topics in a single volume will help facilitate the growth and development of experimentation in political science.

Parts of this chapter come from Druckman et al. (2006).

1. The Evolution and Influence of Experiments in Political Science

Social scientists answer questions about social phenomena by constructing theories, deriving hypotheses, and evaluating these hypotheses by empirical or conceptual means. One way to evaluate hypotheses is to intervene deliberately in the social process under investigation. An important class of interventions is experiments. An experiment is a deliberate test of a causal proposition, typically with random assignment to conditions.[1] Investigators design experiments to evaluate the causal impacts of potentially informative explanatory variables.

Although scientists have conducted experiments for hundreds of years, modern experimentation made its debut in the 1920s and 1930s. It was then that, for the first time, social scientists began to use random assignment in order to allocate subjects to control and treatment groups.[2] One can find examples of experiments in political science as early as the 1940s and 1950s. The first experimental paper in the *American Political Science Review* (*APSR*) appeared in 1956 (Eldersveld 1956).[3] In that study, the author randomly assigned potential voters to a control group that received no messages or to treatment groups that received messages encouraging them to vote via personal contact (which included phone calls or personal visits) or via a mailing. The study showed that more voters in the personal contact treatment groups turned out to vote than those in either the control group or the mailing group; that is, personal contact caused a relative increase in turnout. A short time after Eldersveld's study, an active research program using experiments to study international conflict resolution began (e.g., Mahoney and Druckman 1975; Guetzkow and Valadez 1981), and, later, a periodic but now extinct journal, *The Experimental Study of Politics*, began publication (also see Brody and Brownstein 1975).

These examples are best seen as exceptions, however. For much of the discipline's history, experiments remained on the periphery. In his widely cited methodological paper from 1971, Lijphart (1971) states, "The experimental method is the most nearly ideal method for scientific explanation, but unfortunately it can only rarely be used in political science because of practical and ethical impediments" (684). In their oft-used methods text, King, Keohane, and Verba (1994) provide virtually no discussion of experimentation, stating only that experiments are helpful insofar as they "provide a useful model for understanding certain aspects of nonexperimental design" (125).

A major change in the status of experiments in political science occurred during the last decades of the twentieth century. Evidence of the change is visible in Figure 1.1. Figure 1.1 comes from a content analysis of the discipline's widely regarded flagship journal, the *APSR*, and shows a sharp increase, in recent years, in the number of articles using a random assignment experiment. In fact, more than half of the 71 experimental articles that appeared in the *APSR* during its first 103 years were published after 1992. Other signs of the rise of experiments include the many graduate programs now offering courses on experimentation, National Science Foundation support for experimental infrastructure, and the proliferation of survey experiments in both private and publicly supported studies.[4] Experimental approaches

1 This definition implicitly excludes so-called natural experiments, where nature initiates a random process. We discuss natural experiments in the next chapter.

2 Brown and Melamed (1990) explain that "[r]andomization procedures mark the dividing line between classical and modern experimentation and are of great practical benefit to the experimenter" (3).

3 Gosnell's (1926) well-known voter mobilization field study was not strictly an experiment because it did not employ random assignment.

4 The number of experiments has not only grown, but experiments appear to be particularly influential in shaping research agendas. Druckman et al. (2006) compared the citation rates for experimental articles published in the *APSR* (through 2005) with the rates for 1) a random sample of approximately six nonexperimental articles in every *APSR* volume where at least one experimental article appeared, 2) that same random sample narrowed to include only quantitative articles, and 3) the same sample narrowed to two articles on the same substantive topic that appeared in the same year as the experimental article or in the year before it appeared. They report that experimental

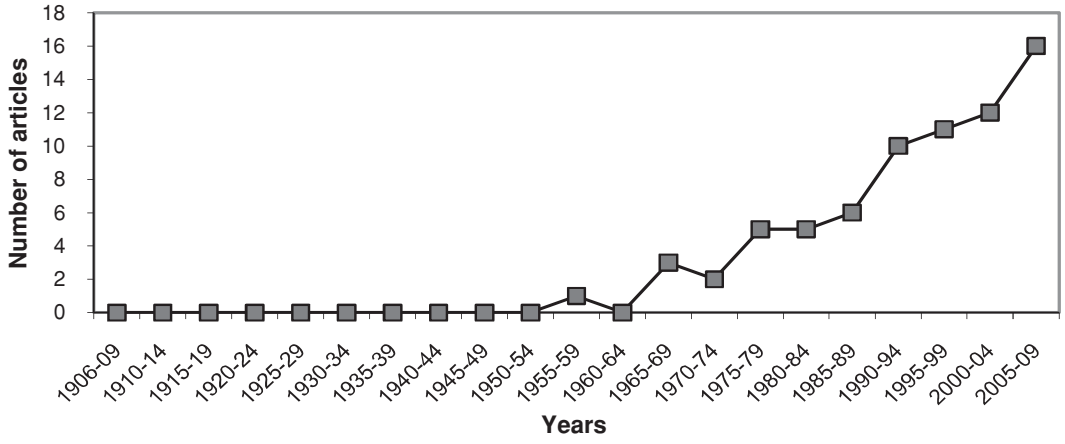

Figure 1.1. Experimental Articles in the *American Political Science Review*

have not been confined to single subfields or approaches. Instead, political scientists have employed experiments across fields and have drawn on and developed a notable range of experimental methods. These sources of diversity make a unifying *Handbook* particularly appealing for the purpose of facilitating coordination and communication across varied projects.

2. Diversity of Applications

Political scientists have implemented experiments for various purposes to address a multitude of issues. Roth (1995) identifies three nonexclusive roles that experiments can play, and a cursory review makes clear that political scientists employ them in all three ways. First, Roth describes "searching for facts," where the goal is to "isolate the cause of some observed regularity, by varying details of the way the experiments were conducted. Such experiments are part of the dialogue that experimenters carry on with one another" (22). These types of experiments often complement observational research (e.g., work not employing random assignment) by arbitrating between conflicting results derived from observational data. "Searching for facts" describes many experimental studies that attempt to estimate the magnitudes of causal parameters, such as the influence of racial attitudes on policy preferences (Gilens 1996) or the price elasticity of demand for public and private goods (Green 1992).

A second role entails "speaking to theorists," where the goal is "to test the predictions [or the assumptions] of well articulated formal theories [or other types of theories]. . . . Such experiments are intended to feed back into the theoretical literature – i.e., they are part of a dialogue between experimenters and theorists" (Roth 1995, 22). The many political science experiments that assess the validity of claims made by formal modelers epitomize this type of correspondence (e.g., Ostrom, Walker, and Gardner 1992; Morton 1993; Fréchette, Kagel, and Lehrer 2003).[5] The third usage is "whispering in the ears of princes," which facilitates "the dialogue between experimenters and policymakers. . . . [The] experimental environment is designed to resemble closely, in certain respects, the naturally occurring environment that is the focus of interest for the policy purposes at hand" (Roth 1995, 22). Cover and Brumberg's (1982) field experiment examining the effects of mail from members of the U.S. Congress on their constituents' opinions

articles are cited significantly more often than each of the comparison groups of articles (e.g., 47%, 74%, and 26% more often, respectively).

5 The theories need not be formal; for example, Lodge and his colleagues have implemented a series of experiments to test psychological theories of information processing (e.g., Lodge et al. 1989; Lodge, Steenbergen, and Brau 1995).

exemplifies an experiment that whispers in the ears of legislative "princes."

Although political scientists might share rationales for experimentation with other scientists, their attention to focal aspects of politically relevant contexts distinguishes their efforts. This distinction parallels the use of other modes of inference by political scientists. As Druckman and Lupia (2006) argue, "[c]ontext, not methodology, is what unites our discipline. . . . Political science is united by the desire to understand, explain, and predict important aspects of contexts where individual and collective actions are intimately and continuously bound" (109). The environment in which an experiment takes place is thus of particular importance to political scientists.

And, although it might surprise some, political scientists have implemented experiments in a wide range of contexts. Examples can be found in every subfield. Applications to American politics include not only topics such as media effects (e.g., Iyengar and Kinder 1987), mobilization (e.g., Gerber and Green 2000), and voting (e.g., Lodge, McGraw, and Stroh 1989), but also studies of congressional and bureaucratic rules (e.g., Eavey and Miller 1984; Miller, Hammond, and Kile 1996). The field of international relations, in some ways, lays claim to one of the longest ongoing experimental traditions with its many studies of foreign policy decision making (e.g., Geva and Mintz 1997) and international negotiations (e.g., Druckman 1994). Related work in comparative politics explores coalition bargaining (e.g., Riker 1967; Fréchette et al. 2003) and electoral systems (e.g., Morton and Williams 1999), and recently, scholars have turned to experiments to study democratization and development (Wantchekon 2003), culture (Henrich et al. 2004), and identity (e.g., Sniderman, Hagendoorn, and Prior 2004; Habyarimana et al. 2007). Political theory studies include explorations into justice (Frohlich and Oppenheimer 1992) and deliberation (Simon and Sulkin 2001).

Political scientists employ experiments across subfields and for a range of purposes. At the same time, many scholars remain unaware of this range of activity, which limits the extent to which experimental political scientists have learned from one another. For example, scholars studying coalition formation and international negotiations experimentally can benefit from talking to one another, yet there is little sign of engagement between the respective contributors to these literatures. Similarly, there are few signs of collaboration among experimental scholars who study different kinds of decision making (e.g., foreign policy decision making and voting decisions). Of equal importance, scholars within specific fields who have not used experiments may be unaware of when and how experiments can be effective. A goal of this *Handbook* is to provide interested scholars with an efficient and effective way to learn about a broad range of experimental applications, how these applications complement and supplement nonexperimental work, and the opportunities and challenges inherent in each type of application.

3. Diversity of Experimental Methods

The most apparent source of variation in political science experiments is where they are conducted. To date, most experiments have been implemented in one of three contexts: laboratories, surveys, and the field. These types of experiments differ in terms of where participants receive the stimuli (e.g., messages encouraging them to vote), with that exposure taking place, respectively, in a controlled setting; in the course of a phone, in-person, or web-based survey; or in a naturally occurring setting such as the voter's home (e.g., in the course of everyday life, and often without the participants' knowledge).[6]

Each type of experiment presents methodological challenges. For example, scholars have long bemoaned the artificial settings of campus-based laboratory experiments and the widespread use of student-aged subjects. Although experimentalists from other

6 In some cases, whether an experiment is one type or another is ambiguous (e.g., a web survey administered in a classroom); the distinctions can be amorphous.

disciplines have examined implications of running experiments "on campus," this literature is not often cited by political scientists (e.g., Dipboye and Flanagan 1979; Kardes 1996; Kühberger 1998; Levitt and List 2007). Some political scientists claim that the problems of campus-based experiments can be overcome by conducting experiments on representative samples. This may be true; however, the conditions under which such changes produce more valid results have not been broadly examined (see, e.g., Greenberg 1987).[7]

Survey experiments, although not relying on campus-based "convenience samples," also raise questions about external validity. Many survey experiments, for example, expose subjects to phenomena they might have also encountered prior to participating in an experiment, which can complicate causal inference (Gaines, Kuklinski, and Quirk 2007).

Field experiments are seen as a way to overcome the artificiality of other types of experiments. In the field, however, there can be less control over what experimental stimuli subjects observe. It may also be more difficult to get people to participate due to an inability to recruit subjects or to subjects' unwillingness to participate as instructed once they are recruited.

Besides where they are conducted, another source of diversity in political science experiments is the extent to which they follow experimental norms in neighboring disciplines, such as psychology and economics. This diversity is notable because psychological and economic approaches to experimentation differ from each other. For example, where psychological experiments often include some form of deception, economists consider it taboo. Psychologists rarely pay subjects for specific actions they undertake during an experiment. Economists, in contrast, often require such payments (Smith 1976). Indeed, the inaugural issue of *Experimental Economics* stated that submissions that used deception or did not pay participants for their actions would not be accepted for publication.[8]

For psychologists and economists, differences in experimental traditions reflect differences in their dominant paradigms. Because most political scientists seek first and foremost to inform political science debates, norms about what constitutes a valid experiment in economics or psychology are not always applicable. So, for any kind of experiment, an important question to ask is: which experimental method is appropriate?

The current debate about this question focuses on more than the validity of the inferences that different experimental approaches can produce. Cost is also an issue. Survey and field experiments, for example, can be expensive. Some scholars question whether the added cost of such endeavors (compared to, say, campus-based laboratory experiments) is justifiable. Such debates are leading more scholars to evaluate the conditions under which particular types of experiments are cost effective. With the evolution of these debates has come the question of whether the immediate costs of fielding an experiment are offset by what Green and Gerber (2002) call the "downstream benefits of experimentation." Downstream benefits refer to subsequent outcomes that are set in motion by the original experimental intervention, such as the transmission of effects from one person to another or the formation of habits. In some cases, the downstream benefits of an experiment only become apparent decades afterward.

In sum, the rise of an experimental political science brings both new opportunities for discovery and new questions about the price of experimental knowledge. This *Handbook* is organized to make the broad range of research opportunities more apparent and

7 As Campbell (1969) states, "... had we achieved one, there would be no need to apologize for a successful psychology of college sophomores, or even of Northwestern University coeds, or of Wistar staring white rats" (361).

8 Of the laboratory experiments identified as appearing in the *APSR* through 2005, half employed induced value theory, such that participants received financial rewards contingent on their performance in the experiment. Thirty-one percent of laboratory experiments used deception; no experiments used both induced value and deception.

to help scholars manage the challenges with greater effectiveness and efficiency.

4. The Volume

In concluding his book on the ten most fascinating experiments in the history of science, Johnson (2008) explains that "I've barely finished the book and already I'm second-guessing myself" (158). We find ourselves in an analogous situation. There are many exciting kinds of experimental political science on which we can focus. Although the *Handbook*'s content does not cover all possible topics, we made every effort to represent the broad range of activities that contemporary experimental political science entails. The content of the *Handbook* is as follows.

We begin, in Part I, with a series of chapters that provide an introduction to experimental methods and concepts. These chapters provide detailed discussion of what constitutes an experiment, as well as the key considerations underlying experimental designs (i.e., internal and external validity, student subjects, payment, and deception). Although these chapters do not delve into the details of precise designs and statistical analyses (see, e.g., Keppel and Wickens 2004; Morton and Williams 2010), their purpose is to provide a sufficient base for reading the rest of the *Handbook*. We asked the authors of these chapters not only to review extant knowledge, but also to present arguments that help place the challenges of, and opportunities in, experimental political science in a broader perspective. For example, our chapters regard questions about external validity (i.e., the extent to which one can generalize experimental findings) as encompassing much more than whether a study employs a representative (or, at least, nonstudent) sample. This approach to the chapters yields important lessons about when student-based samples, and other common aspects of experimental designs, are and are not problematic.[9]

Part II contains four essays written by prominent scholars who each played an important role in the development of experimental political science.[10] These essays provide important historical perspectives and relevant biographic information on the development of experimental research agendas. The authors describe the questions they hoped to resolve with experiments and why they believe that their efforts succeeded and failed as they did. These essays also document the role experiments played in the evolution of much broader fields of inquiry.

Parts III to VIII of the *Handbook* explore the role of political science experiments on a range of scholarly endeavors. The chapters in these parts clarify how experiments contribute to scientific and social knowledge of many important kinds of political phenomena. They describe cases in which experiments complement nonexperimental work, as well as cases where experiments advance knowledge in ways that nonexperimental work cannot. Each chapter describes how to think about experimentation on a particular topic and provides advice about how to overcome practical (and, when relevant, ethical) hurdles to design and implementation.

In developing this part of the *Handbook*, we attempted to include topics where experiments have already played a notable role. We devoted less space to "emerging" topics in experimental political science that have great potential to answer important questions but are still in early stages of development. Examples of such work include genetic and neurobiological approaches (e.g., Fowler and Schreiber 2008), nonverbal communication (e.g., Bailenson et al. 2008), emotions (e.g., Druckman and McDermott 2008), cultural norms (e.g., Henrich et al. 2004), corruption

9 Perhaps the most notable topic absent from our introductory chapters is ethics and institutional review boards. We do not include a chapter on ethics because it is our sense that, to date, it has not surfaced as a major issue in political science experimentation. In addition, more general relevant discussions are readily available (e.g., Singer and Levine 2003; Hauck 2008). Also see Halpern (2004) on ethics in clinical trials. Other methodological topics for which we do not have chapters include Internet methodology and quasi-experimental designs.

10 Of course, many others played critical roles in the development of experimental political science, and we take some comfort that most of these others have contributed to other volume chapters.

(e.g., Ferraz and Finan 2008; Malesky and Samphantharak 2008), ethnic identity (e.g., Humphreys, Posner, and Weinstein 2002), and elite responsiveness (e.g., Esterling, Lazer, and Neblo 2009; Richardson and John 2009). Note that the *Handbook* is written in such a way that any of the included chapters can be read and used without having read the chapters that precede them.

The final part of the book, Part IX, covers a number of advanced methodological debates. The chapters in this part address the challenges of making causal inferences in complex settings and over time. As with the preceding methodological chapters, these chapters do more than review basic issues; they also develop arguments on how to recognize and adapt to such challenges in future research.

The future of experimental political science offers many new opportunities for creative scholars. It also presents important challenges. We hope that this *Handbook* makes the challenges more manageable for you and the opportunities easier to seize.

5. Conclusion

In many scientific disciplines, experimental research is the focal form of scholarly activity. In these fields of study, disciplinary norms and great discoveries are indescribable without reference to experimental methods. For the most part, political science is not such a science. Its norms and great discoveries often come from scholars who integrate and blend multiple methods. In a growing number of topical areas, experiments are becoming an increasingly common and important element of a political scientist's methodological tool kit (see also Falk and Heckman 2009). Particularly in recent years, there has been a massive expansion in the number of political scientists who see experiments as useful and, in some cases, transformative.

Experiments appeal to our discipline because of their potential to generate stark and powerful empirical claims. Experiments can expand our abilities to change how critical target audiences think about important phenomena. The experimental method produces new inferential power by inducing researchers to exercise control over the subjects of study, to randomly assign subjects to various conditions, and to carefully record observations. Political scientists who learn how to design and conduct experiments carefully are often rewarded with a clearer view of cause and effect.

Although political scientists disagree about many methodological matters, perhaps there is a consensus that political science best serves the public when its findings give citizens and policymakers a better understanding of their shared environs. When such understandings require stark and powerful claims about cause and effect, the discipline should encourage experimental methods. When designed in a way that target audiences find relevant, experiments can enlighten, inform, and transform critical aspects of societal organization.

References

Bailenson, Jeremy N., Shanto Iyengar, Nick Yee, and Nathan A. Collins. 2008. "Facial Similarity between Voters and Candidates Causes Influence." *Public Opinion Quarterly* 72: 935–61.

Brody, Richard A., and Charles N. Brownstein. 1975. "Experimentation and Simulation." In *Handbook of Political Science* 7, eds. Fred I. Greenstein and Nelson Polsby. Reading, MA: Addison-Wesley, 211–63.

Brown, Stephen R., and Lawrence E. Melamed. 1990. *Experimental Design and Analysis*. Newbury Park, CA: Sage.

Campbell, Donald T. 1969. "Prospective: Artifact and Control." In *Artifact in Behavioral Research*, eds. Robert Rosenthal and Robert Rosnow. New York: Academic Press, 264–86.

Cover, Albert D., and Bruce S. Brumberg. 1982. "Baby Books and Ballots: The Impact of Congressional Mail on Constituent Opinion." *American Political Science Review* 76: 347–59.

Dipboye, Robert L., and Michael F. Flanagan. 1979. "Research Settings in Industrial and Organizational Psychology: Are Findings in the Field More Generalizable Than in the Laboratory?" *American Psychologist* 34: 141–50.

Druckman, Daniel. 1994. "Determinants of Compromising Behavior in Negotiation: A Meta-Analysis." *Journal of Conflict Resolution* 38: 507–56.

Druckman, James N., Donald P. Green, James H. Kuklinski, and Arthur Lupia. 2006. "The Growth and Development of Experimental Research Political Science." *American Political Science Review* 100: 627–36.

Druckman, James N., and Arthur Lupia. 2006. "Mind, Will, and Choice." In *The Oxford Handbook on Contextual Political Analysis*, eds. Charles Tilly and Robert E. Goodin. Oxford: Oxford University Press, 97–113.

Druckman, James N., and Rose McDermott. 2008. "Emotion and the Framing of Risky Choice." *Political Behavior* 30: 297–321.

Eavey, Cheryl L., and Gary J. Miller. 1984. "Bureaucratic Agenda Control: Imposition or Bargaining?" *American Political Science Review* 78: 719–33.

Eldersveld, Samuel J. 1956. "Experimental Propaganda Techniques and Voting Behavior." *American Political Science Review* 50: 154–65.

Esterling, Kevin Michael, David Lazer, and Michael Neblo. 2009. "Means, Motive, and Opportunity in Becoming Informed about Politics: A Deliberative Field Experiment Involving Members of Congress and Their Constituents." Unpublished paper, University of California, Riverside.

Falk, Armin, and James J. Heckman. 2009. "Lab Experiments Are a Major Source of Knowledge in the Social Sciences." *Science* 326: 535–38.

Ferraz, Claudio, and Frederico Finan. 2008. "Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes." *Quarterly Journal of Economics* 123: 703–45.

Fowler, James H., and Darren Schreiber. 2008. "Biology, Politics, and the Emerging Science of Human Nature." *Science* 322: 912–14.

Fréchette, Guillaume, John H. Kagel, and Steven F. Lehrer. 2003. "Bargaining in Legislatures." *American Political Science Review* 97: 221–32.

Frohlich, Norman, and Joe A. Oppenheimer. 1992. *Choosing Justice: An Experimental Approach to Ethical Theory*. Berkeley: University of California Press.

Gaines, Brian J., James H. Kuklinski, and Paul J. Quirk. 2007. "The Logic of the Survey Experiment Reexamined." *Political Analysis* 15: 1–20.

Gerber, Alan S., and Donald P. Green. 2000. "The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout." *American Political Science Review* 94: 653–63.

Geva, Nehemia, and Alex Mintz. 1997. *Decision-Making on War and Peace: The Cognitive-Rational Debate*. Boulder, CO: Lynne Rienner.

Gilens, Martin. 1996. "'Race Coding' and White Opposition to Welfare." *American Political Science Review* 90: 593–604.

Gosnell, Harold F. 1926. "An Experiment in the Stimulation of Voting." *American Political Science Review* 20: 869–74.

Green, Donald P. 1992. "The Price Elasticity of Mass Preferences." *American Political Science Review* 86: 128–48.

Green, Donald P., and Alan S. Gerber. 2002. "The Downstream Benefits of Experimentation." *Political Analysis* 10: 394–402.

Greenberg, Jerald. 1987. "The College Sophomore as Guinea Pig: Setting the Record Straight." *Academy of Management Review* 12: 157–59.

Guetzkow, Harold, and Joseph J. Valadez, eds. 1981. *Simulated International Processes*. Beverly Hills, CA: Sage.

Habyarimana, James, Macartan Humphreys, Daniel Posner, and Jeremy M. Weinstein. 2007. "Why Does Ethnic Diversity Undermine Public Goods Provision?" *American Political Science Review* 101: 709–25.

Halpern, Sydney A. 2004. *Lesser Harms: The Morality of Risk in Medical Research*. Chicago: The University of Chicago Press.

Hauck, Robert J.-P. 2008. "Protecting Human Research Participants, IRBs, and Political Science Redux: Editor's Introduction." *PS: Political Science & Politics* 41: 475–76.

Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, and Herbert Gintis, eds. 2004. *Foundations of Human Society*. Oxford: Oxford University Press.

Humphreys, Macartan, Daniel N. Posner, and Jeremy M. Weinstein. 2002. "Ethnic Identity, Collective Action, and Conflict: An Experimental Approach." Paper presented at the annual meeting of the American Political Science Association, Boston.

Iyengar, Shanto, and Donald R. Kinder. 1987. *News That Matters: Television and American Opinion*. Chicago: The University of Chicago Press.

Johnson, George. 2008. *The Ten Most Beautiful Experiments*. New York: Alfred A. Knopf.

Kardes, Frank R. 1996. "In Defense of Experimental Consumer Psychology." *Journal of Consumer Psychology* 5: 279–96.

Keppel, Geoffrey, and Thomas D. Wickens. 2004. *Design and Analysis: A Researcher's Handbook*.