The Art of Empirical Investigation

Julian L. Simon

With a new introduction by James E. Katz

Research

Observer Bias

Sampling Error

Useful Knowledge
The Art of
EMPIRICAL
INVESTIGATION
The Art of
EMPIRICAL
INVESTIGATION

Julian L. Simon

With a new introduction by
James E. Katz
Introduction to the Transaction Edition  xiii
Preface to Second Edition  xix
Acknowledgments to Second Edition  xxii
Acknowledgments to First Edition  xxiii
Prologue  xxiv

1. Introduction  3
   1. Purpose of the book  3
   2. What kinds of research are called “empirical”?  5
   3. The place of statistics in the study of research methods  9
   4. Some general remarks  10

PART ONE: The Process of Social-Science Research

2. The Language of Research: Definitions and Validity  12
   1. Operational definitions  12
   2. Numbers and operational definitions  17
   3. Validity and reliability  21
   4. Writing scientific reports  23
   5. Communicating with yourself  25
   6. Honesty in scientific reporting  25
   7. Summary  28

3. Basic Concepts of Research  30
   1. Variables in general  30
   2. The dependent variable  31
   3. The independent variable  31
   4. Parameters  31
   5. The functional form  33
   6. Assumption, theory, deduction, hypothesis, fact, and law  34
   7. Universe and sample  36
   8. The ideal causal-study design  36
   9. Ceteris paribus  39
   10. Summary  47

4. Types of Empirical Research  43
   1. Introduction  43
   2. Case-study descriptive research  44
   3. Classification research  46
   4. Measurement and estimation  49
   5. Comparison problems  50
   6. Research that tries to find relationships  52
3. Advantages of experiments  153
4. Disadvantages of experiments  156
5. Summary  158

11. Designing Experiments  160
1. Simple uncontrolled exploratory experiments  160
2. Before-and-after experiments  162
3. Matched-groups design  163
4. Simple randomized-groups design  164
5. Some additional one-variable designs  166
6. Multivariate designs  167
7. Experimental designs to study delayed effects  171
8. Summary  171

1. Time series—the long view  175
2. The cross section—the wide view  176
3. Causes of differences in results from time-series and cross-sectional studies  182
4. Designs for studying changes over time  183
5. The panel  185
6. Summary  187

13. Surveys: Pro, Con, and How to Do Them  190
1. The nature of surveys  190
2. Advantages of the survey method for relationship research  191
3. Disadvantages of the survey method for relationship research  192
4. Descriptive surveys  193
5. The steps in executing a survey  200
6. Summary  201

14. Some Other Qualitative and Quantitative Techniques  203
1. Deductive reasoning  203
2. The case study  206
3. Participant observation  207
4. Expert opinion  208
5. Content analysis  211
6. Simulation  215
7. Summary  217

15. Classifying, Measuring, and Scaling  219
1. Classifying  220
2. Measuring and scaling  227
3. Strengths of scales  230
4. Summary  233

16. Scaling Human Responses  235
1. Types of mental activity to be measured  236
2. Stimulus and response scales  238
3. Simple composite scales  239
4. Choice of scales  240
5. Summary  243
Contents

22. Obstacles to the Study of Changes Over Time 336
   1. Summary 340

23. Obstacles to the Search for Causal Relationships 341
   1. Causation by a "hidden third factor" 342
   2. Multivariable causation 344
   3. Confounding the dependent and independent variables 348
   4. Feedback; interaction between dependent and independent variables 351
   5. Summary 352

24. The Master Obstacle: Cost 354
   1. Procedure to produce data for study 355
   2. Sampling and cost 356
   3. Sample size 356
   4. How many variables to investigate 357
   5. Amount of tolerable bias 358
   6. Managerial cost efficiency in research 358
   7. Summary 360

PART FOUR: Extracting the Meaning of Data

25. Analysis of Simple Data: Association and Regression 362
   1. Testing for the existence of simple relationships 363
   2. Characterizing the form of an association 366
   3. Regression analysis: characterizing a causal relationship 369
   4. Evaluating the strength and importance of observed differences and relationships 373
   5. Summary 376

   1. Searching for new relationships within the data 378
   2. Reading the meanings of data 382
   3. Refining observed associations into causal relationships 385
   4. Generalizing and predicting 390
   5. The tension between facts and theory or hypothesis 394
   6. Summary 395

27. Inferential Statistics: Introduction 397
   1. Types of statistics 397
   2. The nature and meaning of probability 398
   3. The concept of independent events 402
   4. The distinction between "probability theory" and "inferential statistics" 403
   5. Translating scientific questions into probabilistic and statistical questions 404
   6. Summary 409

28. Probability and Hypothesis Testing by the Monte Carlo Method 410
   1. Introductory problems 410
   2. The general procedure 416
   3. Hypothesis testing 418
   4. Summary 429
Contents

29. Hypothesis Testing with Measured Data 430

30. Correlation and Other Statistical Issues 437
   1. Correlation and Association 437
   2. Advantages and disadvantages of the Monte Carlo method 448
   3. The multiplication rule 449
   4. Evaluating the significance of "dredged-up" relationships 450
   5. Summary 456

31. How Big a Sample? 459
   1. Samples in descriptive research 461
   2. Step-wise sample-size determination 469
   3. Summary 470

32. The Concept of Causality in Social Science, with Notes on Prediction, Law, Explanation, and Function 472
   1. Causality 472
   2. The problem of definition 477
   3. Generalization about the meaning of "cause and effect" 485
   4. Postscript: time, antecedence, and causality 494
   5. Summary 497

PART FIVE: Epilogue

33. The Nature, Powers, and Limits of Social Science 500
   1. What makes an investigation scientific? 500
   2. What can empirical research do? 502
   3. The necessity of knowing the purpose of the research 505
   4. The special powers and limitations of social-science research 507
   5. The ethics of research 510
   6. The present and future of social-science research 512

Bibliography 514
Index 539
I first encountered Julian Simon’s book on research methods in 1969, shortly after it was published. The country and campus were in the midst of an intense war in Southeast Asia and on the streets of America. It was a war not only for the hearts and minds of distant peoples in a twilight struggle, but for the heart and soul of America. Now, more than thirty years later, the United States is involved in another dramatic struggle, both in the twilight regions of the world and in the full daylight here at home. Social science research was part of the domestic and international war effort then. Today it is even more involved in both fronts, albeit in some dramatically different ways.

Then, I was a callow undergraduate; sociology was my game and societal improvement was my aim. I was attending a large Midwestern university that could plausibly have served as the model for the one in Sinclair Lewis’ Arrowsmith. In those years, it seemed that the discipline of sociology offered a good way to take science and apply it to improving the state of the world. Or, to draw upon phrasing that might reflect my study of Aristotle, I sought to take science, the passion of my head, and harness it to service, the passion of my hand. In that era, during the last fading echoes of an ethos concerned about standards, all sociology majors at my university were required to take statistics, two years of foreign language, a year of English composition, advanced math, public speaking, civics, and physical education. And research methods.

Unlike some of my pals, I was enthusiastic about the prospect of taking this course bearing what they insisted was a soporific title. My reasoning was that I had been unhappy with the squishy nature of what in the sophomore level courses had paraded as knowledge. I imagined that with the modern tools offered by quasi-experimental designs, operations
management, and computer technology, the situation could be rectified. Although naive in retrospect, my thinking at the time was that as, with improved instrumentation and careful observation, astronomy had superceded (but sadly never displaced) astrology, so too some new discipline (socionomy?) would, by exploiting better research methods, supercede sociology.

So I wanted to get started on using the tools of science—I was sure they existed on some socially isolated Platonic level—to discover the equally real and unvarying laws by which social behavior and organizational structures are governed. But then came the rub: how could one uncover possible candidates, and then select the genuine article? How could they be applied so that they could perform at their best? Certainly those were difficult questions, but equally certain was that a research methods course would be the place to start. And to take the course one would need to get the assigned textbook.

Thus it was that a few weeks before the new semester began, I found myself heading to the university bookstore to buy the required text. To get to the bookstore one had to first wade through the Pow-Wow Room, a garish, noisy, and insensitively named cafeteria that was the hub of campus social and intellectual life from 11:30 to 2:00. By passing into the campus bookstore through its paired Safe-T-Glass doors, a hermetic seal was made, stifling the blare of the Strawberry Alarm Clock or Jimi Hendrix. Simply by dropping one’s bag at the security desk and pushing through a clanking turnstile, one would pass from a world of free speech sit-ins and agi-prop theater happenings into an intimately planned, quiescent commercial environment. Low-fidelity ceiling speakers filled the air with the Hollywood Strings’ rendition of the Beatles’ “You say you want a revolution.” Students turned stock boys for a semester hummed along as they stuffed racks with cards showing Walter Keane’s giant orbied waifs shyly kissing each other.

Strolling to the textbook section, I passed beneath sagging Peter Max Love posters and mawkish Erich Segal Love Story posters. I then threaded between Babel towers of compact cardboard sample boxes—blush-pink paisley boxes for girls, light-blue paisley boxes for boys—filled with sun-dry freebie personal care products that declared themselves to be “revolutionary.” These, I confess, had their intended effect, exercising a powerful lure. Though feeling a bit like Ulysses tied to the old mast, I nonetheless pressed on to my goal: at last the textbook section hove into view.

In contrast to the intensely colorful commercial extravaganza, the textbook zone was a thicket of Army surplus metal utility shelves. One had to step cautiously for columns of textbooks had tumbled out into the narrow aisles. Despite a few blind alleys, and many pauses to consider what students in other courses would be reading, it was but a short time until I located where the required textbook for the research methods course had been left. Unlike Intro Psych, Intro Soc and Intro Poli Sci texts, this research methods book was neither colorful in design nor revo-
lutionary in its graphics. It was also unlike its slender compatriots that addressed research methods of other social science disciplines. *Ecce libris!* This volume had a girth that made it seem like a millstone.

I pulled a copy from the shelf and began leafing. Surpassing 500 pages, the book's impression of intellectual and physical weightiness was reinforced by its small font. Worse, the font had an exaggerated serif that suggested it had been set in italics. No photos were provided for visual interest nor were there any colophons to attract the eye. Still worse: it claimed it was only a basic introduction to the subject! Perhaps, I thought to myself, the saturnine judgments of my friends had been correct.

But then I started to read the text, jumping here and there, and my spirits rose. It seemed to be written by someone who cared about me, not in a condescending way, but in a way that would allow me to help myself reach my goals. The examples drew upon every aspect of, well, life and one's concerns in the public and private spheres. Beginning at that moment under those fluorescent lights, next to those gray surplus shelves, I began to develop a bond of affection for the author and, more importantly, for what seemed to be his mission.

In this book, as would be the case with many of his later articles and interviews, Julian Simon achieves the voice of a droll, world-wise uncle. Using straightforward terms and charming homilies, he presents his and other top-flight researchers' insights in a benign, conversational manner. One is reminded of the image of a country doctor, though in this case it is a country doctor of philosophy. Professor Simon covers the vast topic of how to conduct research not only clearly but also reasonably. He is out with the pedantic and in with the effective. He displays charm, warmth, and self-deprecating wit as he pursues his reasoned dialogue, ever informed by the pragmatic in pursuit of the ideal but ever mindful of the inherent shortcomings necessitated by human limitations.

Professor Simon has distilled an immense amount of knowledge about the realities of the “how-to’s” of research without reverting to cookbook-like lists (both the bane and blessing of contemporary texts). These insights have been gained, it becomes clear, through hands-on experiences (and, also equally plain, some of them rather grim). Gears are shifted rapidly at times, in the interest of making points clearly. Hence, one will find snippets of uxorial chitchat placed near questions of macro-economic modeling. Ultimately Professor Simon, it is clear, cared passionately about the subject, and justifies his firm belief that reason and evidence is the best hope for mankind's continued survival and welfare.

Within a few years of this first encounter I was in hot pursuit of an advanced degree and was fortunate enough to have a turn to helm my own research methods course. I knew with the perfect certainty of one possessed which textbook I wanted for my class, and you, dear reader, can surely guess what it was. But I was astounded to learn that the book had gone out of print. “How could they?” I asked myself. A feeling of betrayal seized me: so quickly this treasure had vanished from the intel-
lectual landscape. I still had my own copy of the book (and which I have kept until this very day). But what about all the future students of research methods? Who would guide and take care of them, I wondered.

In the three long decades since the book went out of print, I have soldiered on with various textbooks. Many were not bad, some were even quite good. But none, despite their individual merits, could measure up to the high standard of Julian Simon’s textbook. Oh sure, they covered much of the same ground. The steps in doing a survey could be, and were, described by these other textbooks. And yes, they did discuss internal and external reliability of controlled experiments. Yet, these were textbooks, if not written by a committee, massaged and homogenized by one. The reasons why the steps were to be undertaken and the problems that might ensue if they were not followed were presented, of course. But there was no real sense of a person struggling to find reasonable answers in any of these lists and condescending, often politically inspired, examples. Absent entirely were appreciations of the ever-surprising ironies and capriciousness of the human animal. Perhaps, worst of all, there was no sense of a shared culture, no feeling for the norms and folkways of the research argument, and no involvement with the invaluable rules of thumb that must guide social scientists as they work in real-world environments. In a word, the textbooks I found seemed to be written by teachers of research methods rather than by researchers. And they seemed to be written for people who are taking the course to fulfill a requirement rather than a personal mission, certainly not for someone on a mission to become a researcher.

So part of my pain has been that new generations of researchers were effectively denied the deeper understanding of “how we know what is so” that Professor Simon’s able assistance would have provided. That is, he was able to convey the realities, lessons, and limitations of doing research, and cover a vast array of approaches, without the necessity of students having to experience the failures and mistakes that would lead to knowing why things should be done in one way as opposed to another. He focuses invariably on why things are done, so that the student can understand and extend the principles to be effective consumers and creators of social science research. By contrast, most other methods textbooks want to give the rules of procedure as derived from other textbooks, so that students are left with myriad lists that have the same intellectual meaning as a tax form.

Since that first research methods course that I taught in 1973, over two thousand students have been under my tutelage in variously named, but essentially the same, principles of research methods classes. (Since the scientific method itself has remained essentially the same for several centuries, I have no qualms about saying my teaching of it has been essentially the same for decades. Of course, the specific details of scientific research programs and tools change dramatically if one speaks of years; the same could be said about my course if one speaks of semes-
ters.) Though I tried many other textbooks, none have proven themselves equal to that of the good professor's.

Above I have given reasons why this research methods book is outstanding. But let me complete the discussion here. I believe that Professor Simon wrote this book based on vast experience doing research, often tutoring directly research assistants who helped on his projects. When it came time to teach his own course, he immediately saw the problem, which is the same one that I have faced for thirty years. The crop of textbooks does not give the student what the student really needs, which is the reasoning underlying research and the reasonable ways to get it done. Learning to do research is akin to learning any other skill, such as swimming or riding a bicycle. (I thank Howard S. Becker for this analogy.) You could present all the chalkboard lectures you might want, but there is essentially only one way that anybody is going to be able to swim or bike ride, and that is by trying to do it. This book is the next best thing to doing research, and doing it with a convivial senior mentor. And it is an excellent first step to beginning to do research, just as knowing about momentum and resistance are to the sports of swimming and biking.

Through the wisdom, perceptiveness, and sensibility of Professor Irving Louis Horowitz, editorial director of Transaction Publishers, it is the happy fate of this work to be at last in print again. It is also a happy and felicitous event for generations of students who are already aspiring to be researchers to have this canny guide. It will also motivate some students who have no interest in becoming researchers as a vocation to become competent amateur researchers who can help themselves and others by understanding the craft. Certainly many students in my future research methods courses will be able to enjoy the down-to-earth reasonableness of an astute practitioner.

A word is due Hanan Selvin, who was this book's midwife. Hanan was a brilliant methodologist in his own right. He was a kind, dedicated teacher who lived for his craft and relished every minute of it. Sadly, Hanan was struck with a progressive degenerative disease that led early to his blindness. But instead of feeling sorry for himself, he began to study the sociology of blindness. I was a young graduate student in the sociology department of Rutgers when he visited the campus for a guest seminar. His lecture was on the tricks of the trade practiced by blind people to get around in a world built for those with sight. I was fortunate enough to make his acquaintance on this occasion, and he provided me with useful methodological advice about the research that I was doing for my dissertation. Hanan clearly cared enough to spend his time helping others learn about the scientific method and its applications to human behavior. Without his strong intervention and encouragement, this book by Professor Simon might never have been published in the first place. So in an important sense, the re-issuing of this textbook is a testimonial to not only the enduring sense of craft, wisdom, and concern of Professor Simon, but of Professor Selvin as well.
Let me return to what I noticed when I examined this book for the first time, back in the bowels of the campus bookstore, back so many decades ago. Unlike the covers of others strewn nearby, this textbook’s cover offered no promise that its purchase would allow a student to change the world for the better. Yet it should have.

This book offers one of the best stories yet written about how to actually conduct research using real tools in a real world. With these tools, one can gain an understanding of the world, not as one would like it to be, but how it is. This is a necessary first step. Meaningful progress is rarely achieved save through the rationality of science, which itself must be informed by sound research. Sound research is the one best key that is dangling before us. If we grasp and apply effectively this key, it will open for us doors to a world of nurturing social environments, personal autonomy, and great opportunities for pursuit of personal fulfillment. Now that Professor Simon’s book is back in print, more students will be able to understand and do research. They will be given sorely needed guidance on how to use the key of research to create a future of progress, abundance, and civility. This is something he wanted for us as much as we would want it for ourselves.

James E. Katz
December 13, 2002
The aim of this second edition is the same as that of the first: to help you get started doing empirical research. The book has been reorganized considerably to better serve that purpose. The chapters on research decisions and procedures now come early in the book (Part II), followed by chapters on the obstacles to research (Part III). And now there are complete chapters on sampling (Chapter 9), experiments (Chapters 10 and 11), surveys (Chapter 13), and scaling (Chapters 15 and 16), plus independent appendices on questionnaire construction (Chapter 20) and interviewing (Chapter 19); in the first edition this material was scattered throughout the book. There is also a new chapter on the relationship between theory and empirical research (Chapter 5).

The first edition of this book is not intellectually obsolete. The practice of good science has changed little since the first edition—or in the last half century, for that matter; this explains why most of the examples have not been changed from the first edition. But the organization of the second edition should be more convenient for you to read. And over the years I’ve learned how to express some ideas better. In addition, the new material fills some needs that readers of the first edition said needed filling. Last but not least, chapter summaries and other pedagogical devices should ease your studying.

Though this is a practical how-to-do-it book, it aims also to teach the basic concepts in the philosophy of science. In my view this is not a contradiction, because the philosophy of science at its best is also a very practical subject, composed of ideas that clarify the nature and meaning of research and help the researcher better understand how to proceed when faced with difficult choices and decisions.

This book may be deceptively easy to read. I have always hated obscurantism and have been emotionally committed to simplicity, perhaps
because of the coincidence of my name with the nursery rhyme "Simple Simon met a pieman. . . ." But simplicity has its drawbacks, such as: (1) Some students concluded that the first edition was so simple to read that it was insufficiently "challenging." Simplicity has a basic psychological difficulty, as expressed by T. S. Eliot about poetry: "A successful poem must be sufficiently simple so it can be understood, but sufficiently difficult so that it cannot be understood immediately." But I am unwilling to make my writing hard to understand so as to be more appealing. (2) Simplicity of expression may fool you into thinking that the ideas discussed here are simple ones. They are not. (3) As someone has said: Seek simplicity—but distrust it.

On the other hand, the book may seem overly long, containing too many examples. But as someone else once said, short writing makes long reading. So I hope you will not be put off by the book's length.

The writing style is more casual than it is conventional, and it does not always meet editors' standards as to what is "correct." Therefore I am pleased to take full responsibility for the language.

Despite that this is a text, it contains some new scientific ideas. Those ideas I most hope that you will notice as being original contributions are, first, causality as a concept requiring appropriate operational definition (Chapter 32) and second, the Monte Carlo approach to basic statistics (Chapters 27 to 31). Since the first publication of the Monte Carlo methods in this book, there has developed a body of experimental evidence that this radically different way of learning and doing statistics is more effective, easier to understand and more effective than the conventional analytic method (Simon, Shevokas, and Atkinson).

About the sex of pronouns in the book: The first edition used "he" for the unidentified person. That will not pass any longer, and never should have passed. However, I do not wish to clutter up the writing with clumsy "he or she" phrases or to weaken it with annoying circumlocutions. The solution I've chosen is simply to sometimes write "he" and sometimes "she," more or less at random, when I mean an unidentified human being. This practice may seem a bit strange at first, but I trust that it will cause no confusion and will be pleasant and efficient. If this practice seems good to the reader, maybe other writers will adopt it, too.

A personal note: The considerable success of the first edition heartened me in many ways. First—and this may in turn hearten other prospective authors—the skeletal first draft of the manuscript was rejected by fifteen (15) publishers before Hanan Selvin pointed out its potential to Random House. Second, I worried that because the book is not narrowly specialized to a single discipline, each teacher would say that it might be good for others but not for him or her. It turns out, however, that there are a good many instructors who believe in a broad education in research methods; this heartens me because I, too, believe that. Third, I worried that the lack of mathematical notation and "sophisticated" and "rigorous" complexity
would deter instructors from adopting the book; but apparently there are plenty of instructors who are more interested in teaching than in impressing students, and that cheers me, too.

In closing, I am grateful for your attention to my thoughts and writing. I benefited from the suggestions and corrections I received from readers of the first edition, and I shall be glad to receive more on this second edition. Thank you.
Hanan Selvin listened to tapes of many chapters for the second edition, especially the new ones, and gave me a flood of useful and delightful comments, only some of which did I have the strength to exploit. Gideon Keren made helpful comments on Chapter 11. And the following long list of people were kind enough to send me evaluations or remarks about the first edition: Marcus Felson, William Ahlhauser, Wayne W. Snyder, Jason Millman, Robert A. Baker, Oleh Wolowyna, Lawrence G. Smith, Lucy W. Sells, Leigh Marlowe, Siamak Movahedi, Michael A. Baer, Joann S. DeLora, Marc D. Magre, Leroy Gruner, John H. Kramer, Jules W. Delambre, Alan C. Acock, Robert W. Shotoln, David M. Krieger, Michael D. Grimes, L. V. Hayes, Robert C. Smigelski, David G. Pfeiffer, Jamie M. Calderon, D. E. W. Holden, David M. Monsees, Jr., L. Fannin, Luther H. Keller, Linda Brookover, Bourque, David Nasatir, Kenneth R. Rothrock, Louis A. Brown, LeRoy Martinson, Güts Wolff. I am grateful to one and all.

More and more my family and our life together sustains me. Wife Rita and children David, Judith, and Daniel give my days the joy and meaning that enable me to write.

Urbana, Illinois, 1978

J.L.S.
For a hundred and more years ingenious social scientists have faced obstacles to getting the empirical knowledge they sought, devised ways to circumvent the obstacles, and then told others what they learned. I am grateful to them.

Hanan Selvin subjected the penultimate draft of this manuscript to the searching scrutiny that most manuscripts need but few are lucky enough to get. The quantity and quality of his critical comments were an author’s dream, and I have in many places appropriated his thought and word without special note. There would be fewer errors and obscurities in the book if I had followed his advice even more diligently.

James W. Carey gave me many exciting and enjoyable hours discussing some of the fundamental concepts. I also benefited from talks with Howard Maclay. Other friends may also recognize their casual observations in these pages. Allen Holmes read the chapters on statistics very carefully and corrected some errors. Dennis J. Aigner was also good enough to look at those chapters. For useful references and suggestions I am grateful to Stanley Friedman, Lewis Goldberg, James H. Lorie, and Louis Schneider.

My greatest debt is to my wife, Rita. Without her encouragement this book would never have been done.

I have tried to cover the wide intellectual area of all the social sciences, and I hope the reader will bear with me when I depart from the substantive fields that I know best. I will appreciate hearing from anyone who can set me straight on any matter or who can give an instructive or interesting example to illuminate a point.

Jerusalem, 1968

J.L.S.
Cast your mind back 450 years. The Governor of Rodera—a small Asian principality—decided to find out why the country's tax revenues were not greater. Here is a classic problem for research, and much of mathematics and social-science techniques was originally invented to improve tax collection.

The Governor first consulted an adviser who had studied in Europe and had learned something of Aristotelian logic. The adviser reasoned that (a) good citizens willingly pay as much tax as they are able to, (b) the folk of Rodera were very good citizens, and therefore (c) the tax collections could not be higher.

The Governor immediately dismissed this syllogistic thinking as pure nonsense because he disbelieved the premises. Even without special training he was smart enough to understand the weakness of bad logic.

Next, the Governor called in the regional tax collectors and put the problem to them, because they were experts in the tax-collection field. After consultation the consensus among the tax collectors was that the people of the country simply could not afford to pay any more taxes. To support this statement the tax men stated that they had already used every possible technique to extract higher taxes—raising valuations of property, checking on hidden assets, and the like—and even their best efforts could not raise more money.

The Governor understood the tax men's analysis of the situation as a self-seeking attempt to make themselves seem competent and hard-working, and he therefore disregarded this “expert opinion.”

The Governor then concluded reluctantly that he would have to find out the answers for himself. He instructed an aide to bring him a handful of typical citizens. The aide brought in some people who were close at hand. They included a few beggars who had been loitering nearby, plus the aide's brother and two young guards. It was obvious to the Governor that these
people were not representative of the population, and therefore it was useless to question them. To put it in modern terms, the Governor recognized that he had a badly biased sample.

The aide was therefore instructed to bring in some “typical” peasants and townspeople, which he did. The Governor then asked them, “Why don’t you pay more taxes?” The first few answers showed the folly of the question, for what he heard were excuses, complaints, entreaties—everything except what seemed to be sensible answers.

The Governor then confined his questions to factual matters. He asked about each person’s property, his crops, family size, and the amount he paid in taxes. Assuming that the answers were true and that he could have them checked, the Governor thought he was getting somewhere.

The information obtained from such a small handful of people clearly was not enough, however, because the group did not include any rich men, any people from the far provinces, any foreign residents, or any representatives of other important classes of people. Nor could the Governor tell from this sample how many people in each class there were. Therefore, the Governor ordered a nation-wide house-to-house census.

Nowadays the data for a complete nation-wide United States census can be collected in months. But transportation and communications were poorer then, and few literate people could be found to collect information. The census therefore required twelve years.

By the fifth year the Governor grew impatient and decided to experiment with the effect of new pressures to increase tax collections. He ordered that anyone who did not pay half again as much tax as in the previous year would have his cattle confiscated. It turned out that the total taxes paid did not increase. The crops were bad that year, and the Governor could not determine whether the taxes would have been as high or higher than otherwise if the crop had been normal. The trouble was that his experiment was uncontrolled, because he had not kept the old tax system in effect in some areas for comparison—that is, his experimental design was incomplete. And worst of all, many people slaughtered their cattle to avoid confiscation, which ruined the country’s meat supply for several years.

By the time the census data were all in hand at the end of the twelfth year, the country’s situation had changed, and the data collected in early years no longer meant much. (Nowadays we know how to take samples to reduce cost and avoid long time lags during which the picture may change.) Furthermore, after twelve years the tax-census data filled a whole warehouse, and another ten years would have been required to interpret them. The Governor gave up the task in disgust and retired to his harem.

The point of this story is that knowledge is often not easy to come by because there are many obstacles in one’s path. And common sense alone is not enough. But by now social science has accumulated a body of tested experience on how best to overcome the obstacles and acquire empirical knowledge efficiently and safely. This book presents some of this accumulated experience.
basic research methods in social science
1. Purpose of the Book

This book is primarily for students who have never before studied or done empirical social-scientific research. I hope that the book contains good advice that will help you get your first research project off the ground successfully and increase your efficiency in later work. As for those of you who will not do empirical research, the book may teach you to distinguish good research from poor research and help you to understand why empirical researchers do things as they do them.

People who have had some training or experience in empirical research may also gain from the elementary level of the discussion. Basic concepts often are bypassed as one rushes to learn the methods of particular fields. Coming back to fundamentals can widen the perspective of an advanced student and fill holes in his knowledge. If an advanced student is to gain something from this book, however, he must have the wisdom to realize that the apparent simplicity of the basic concepts is often deceiving. For example, everyone knows that ceteris paribus—holding "all other things equal"—is important. But the more research you do, the more you realize how complex is the ceteris paribus idea, how difficult it is to choose the right ceteris paribus conditions, and how often research is useless because other things really were not made sufficiently equal.

The book is intended for future producers of research, of course. But many people who study research methods will never produce research; rather, they will be research consumers, in their jobs and as citizens. For this important latter group of people, the aim of the book is to teach how to
evaluate research done by others—to know which research is good and which is not; where the weak spots are in a piece of research and how important they are; and whether a research finding, whether presented in the professional literature, in an informal report, or in the popular press, is likely to be valid or not. For example, Volvo recently advertised that “90 percent of the Volvos sold in the last eleven years in the U.S. are still on the road.” A good course in research methods is likely to alter your impression of what that claim might mean.

I hope that you yourself carry out some empirical research—no matter how small in scope—as you read this book. It is not enough to study empirical research the way one studies astronomy, economics, psychology, and other academic subjects. Reading about research principles is certainly useful. But research is not entirely an academic subject. Rather, it is largely an art, a how-to-do-it subject like musical composition, writing advertisements, or swimming.¹ You never really know how to do research until you do it, any more than you can know how to swim after only pool-side instruction. You must jump into the water, thrash around, and gradually improve with practice. For the same reason, skill in empirical research requires experience—that is, doing research of your own and criticizing the research of others.² For example, when you do a piece of research you cannot fail to learn just how complicated even the simplest research really is.

Eventually you would learn much that is in this book by trial and error. Instruction can only hasten the process and make it less painful by showing what has and has not worked for others. But that is as far as the teaching can go. (Reading what others say about research can be enormously profitable, however, if you will benefit by the experience of others.)

Now a theme that will recur throughout the book: There is never a single, standard, correct method of carrying out a piece of research. Do not wait to start your research until you find out the proper approach, because there are always many ways to tackle a problem—some good, some bad, but probably several good ways. There is no single perfect design. A research method for a given problem is not like the solution to a problem in algebra. It is more like a recipe for beef stroganoff; there is no one best recipe.

For technical matters, too, there may be several satisfactory techniques, and there is no cut-and-dried answer. For example, if you want to do a

---

¹ A great mathematician applies the same analogy to his sort of work, too: “Solving [mathematical] problems is a practical skill like, let us say, swimming. We acquire any practical skill by imitation and practice” (Polya, p. 4).

² Also valuable are the too-rare accounts of how and why a piece of research was done. Three short descriptions of this sort can be found in W. Wallis and H. Roberts (Chap. 3). Anthropologists often weave such material into their writings; W. Whyte (2nd ed., Appendix A) has a delightful and useful description of this sort. K. Colby’s account of Semmelweis and childbed fever is excellent (pp. 44–50). R. Braidwood’s “Biography of a Research Project” is a fascinating description of archaeological research. P. Hammond offers some excellent examples. C. Mills (pp. 199 ff.) interweaves a description of one of his labors with marvelous general advice on craftsmanship and how to get ideas and do social analysis. J. Madge has written a book full of such accounts of sociological research.
questionnaire survey, should you interview by mail, by telephone, or in person? Chapter 13 discusses the pros and cons of each method, but eventually sound judgment is required for this technical decision; no handy rule book can make such decisions for you. Or, should you pay your subjects in an experiment or survey? Again there is no pat answer; instead, this book tries to give you some principles that will help you to make sound technical decisions about your research methods.

In this book the word “method” refers to empirical techniques and devices of various sorts. To the philosophers the term refers to “scientific method”—the whole process of getting knowledge, including the theoretical and the empirical steps. But even then it is surely true that there is no single “scientific method.” “... There is no one scientific method ... there will be as many different scientific methods as there are fundamentally different kinds of problems” (Northrop, pp. ix, 19). “The scientific method, as far as it is a method, is nothing more than doing one’s damnedest with one’s mind, no holds barred” (Bridgman, p. 450).

2. What Kinds of Research Are Called “Empirical”?

This book deals with empirical research and not with scientific speculation. Much of scientific work consists of thinking up ideas about the nature of the world, generalizing from observed facts to scientific “laws,” and developing logical systems that are called “theories” or “models.” But simply as a division of labor the speculative part of science is not covered here. The subject of this book is “getting the facts.” Of course empirical work goes beyond “mere” observation and description, and it is inextricably intertwined with explaining nature and making predictions about it. But the process of thinking up explanations or hypotheses about nature and its laws is beyond our scope here.4

Nor is this book concerned with the logical process of building scientific theories. That is, it does not deal with the process of finding the logical relationships between various scientific statements or with the process of developing generalizations or scientific laws. Rather, its subject is the less glamorous craft of producing and examining factual and material evidence and sense data to develop the descriptions, measurements, comparisons, and tests of hypothesized relationships that are themselves part of the speculative side of scientific work.

Lest there be misunderstanding, I emphasize that a good idea is the

3. “... [T]he adjective empirical, in its combinations with various nouns, appears to denote observations and propositions primarily based on sense experience and/or derived from such experience by methods of inductive logic, including mathematics and statistics” (The Dictionary of the Social Sciences, p. 237). The crucial distinction here is between empirical research and theorizing, though the two activities are very much interdependent.

4. The elusive phenomenon of scientific discovery is discussed at length by many writers. See, for example, J. Young, A. Bachrach (Chap. 1), W. Beveridge (Chaps. 5, 6), C. Mills (Appendix, especially pp. 200–201, 211–217).
keystone of an empirical study. Mere data collection and measurement are worthless unless the subject is important. Theory is often the fount of important ideas for empirical research, and sound theory is of inestimable value in any field. The relationship between theory and empirical research is explored in Chapter 5.

Finally, “empirical research” excludes knowledge obtained by consulting authorities, in books or in person. It includes only knowledge obtained from data resulting from first-hand observations, either by you or by someone else. Reexamination of data collected by others, such as U.S. Census data, is empirical research, of course.

Most of the examples in this book are drawn from “pure” research. But I also draw examples concerning policy decisions from the “applied” social sciences, and they often have a dollars-and-cents orientation. Applied research methods are sometimes more sophisticated than are methods used in pure research (Stouffer, 1950a, pp. 198–9), because it is possible to work up some calculation, even though crude, to compare the benefits expected from a research method against the cost of doing the research with that method. Such calculation leads to efficiency in research. (Pure research can be defined as research whose social or economic payoff is far in the future, whereas applied research is expected to have a quick payoff. But pure research is often done without any thought at all about payoff, just to satisfy the desire to understand.)

Here are two examples of how dollars-and-cents calculation in an applied problem helps to make sensible decisions about research: First, an advertiser can calculate whether comparing two advertisements in a split-run test is worth the cost of the test. One can also compare the costs and benefits of a split run against the costs and benefits of other types of advertising research. Second, a candy firm can sensibly calculate its sample size when comparing two new flavors of candy. It can reckon the relevant costs and dollar benefits. It is much harder to determine the sample size sensibly when I. Pavlov, for example, studies how the flow of saliva in dogs can be conditioned to the sound of a bell. The possible benefits of such pure research—some intangible gain in the quality of human life perhaps far in the future or merely the satisfaction of an urge to understand our world better—are less predictable and much more difficult to evaluate in money terms to balance against the money cost of the research. Yet the decisions must be made anyway.

Our examples come from the social sciences, not only because the book is

5. But see the section on “expert opinion” in Chapter 14.
6. In a split-run test a magazine or newspaper arranges to print two advertisements so that each appears automatically in every other copy in each stack of copies that is sent out. This is the closest thing to a perfect experiment.
7. It is always conceptually possible to develop some rational calculation of the value of a piece of research and its various outcomes so as to have some rational guidelines for research decisions (see Chapter 8; Wilson 86–87; Schlaifer. But in most pure research such calculation is not done because it is so difficult to do meaningfully.
intended for students of the social sciences, but also because empirical scientific methods have been used with greater variety and greater subtlety in the social sciences than in the natural sciences (Chapter 33 defends this claim).

We shall turn frequently to a few such famous studies as Alfred Kinsey's research on sexual behavior, employment and unemployment surveys, Hermann Ebbinghaus' learning experiments, Presidential election polls, television-audience ratings, Sigmund Freud's case history of Anna O., the U.S. Surgeon General's Report on Smoking and Health, and Ivan Pavlov's work on conditioning reflexes. These studies have been chosen for several reasons: They are inherently interesting; most students have a general knowledge of them from survey courses or general reading; most of the studies have been repeated or scrutinized closely by outside experts (for example, the American Statistical Association appointed three top statisticians to report on Kinsey's methods); and they show us a wide variety of methods and a broad representation from the social sciences.

I also refer frequently to my own work, despite its limitations, because I know exactly what went into the work—the difficulties, decisions, errors, corrections of errors, and the order in which things took place. One cannot have such intimate knowledge of anyone else's work. Yet it is these details and decisions, seldom written about, that are hardest for the novice to understand and master and that one usually learns only by serving as an apprentice (that is, "graduate assistant") or by trial and error in one's own work.

The book emphasizes the design and plan of research, rather than the analysis of research data. Except for studies that reanalyze data collected by others, the most important and interesting decisions arise at the design stage, or at least they should arise then. If you postpone these decisions until after the data have been collected, you may suffer heartbreak and wasted expense.

Here are three brief examples of the importance of good design: First, two library scientists sought to determine the proportions of various-sized books in research libraries. So they measured the heights of each of the hundreds of thousands of books in a major library. With a little planning and sound design they would have needed only to measure a fraction of that number of books. And with sound design the results could also have been applied to libraries other than the one they studied.

Second, a family-planning group tested one birth-control propaganda campaign in Village A against another propaganda campaign in Village B, forgetting that subsequent differences in birth rates and contraception-acceptance rates might reflect basic differences between the two villages unconnected to the campaigns, rather than only the differences between the campaigns. Sound planning would complete the design by alternating both campaigns in the two villages or by other methods. The same sort of error has often been made by experimental psychologists and sociologists.
Third, our understanding of voting behavior has been greatly enhanced by the use of the panel method, in which the same people are quizzed about their voting behavior several times during the same election campaign. These repeated observations make it possible to understand the mechanism of voting and vote shifting in ways that are impossible without the panel design.

Mistakes at the design stage can be mended only at great extra cost, or not at all. By comparison, mistakes at the analysis stage can be remedied at slight or no cost as long as the mistakes have not gotten into print or been acted upon.

Think through the research design carefully in advance. Failing to consider all the necessary details at the design stage because of procrastination or mental laziness is one reason that many researchers get very little done. Of course not everything can be foreseen, especially in exploratory studies, but one should use as much foresight as possible. It is useful to talk to your friends about the design, and to prepare an outline of it; both processes reveal fuzziness in your design thinking.

The design of a piece of research must depend upon the particular purpose that the research is intended to serve; this is a message I shall repeat again and again. For example, the Internal Revenue Service publishes statistics on the amount of advertising done by groups of firms that sell various products, and these statistics include all firms' advertising, because the government requires data on the entire economy. But the statistics gathered and published by industry trade associations cover only the leading firms, because the industry-collected statistics are designed to meet only the information needs of the industry. For another example, a psychologist studying the relative memorability of beginning, middle, and ending portions of messages will use a different test of memory (perhaps a list of nonsense syllables) than will a psychologist who is studying how many pieces of information a radio operator can remember accurately. You must ask yourself repeatedly: Exactly what do I want to find out? and why? If you can answer these questions clearly and precisely, you have gone a long way toward creating a satisfactory research design.

Part One begins with Chapter 2 on the language of science, which is inseparable from science itself. Chapter 3 discusses basic concepts such as variable, function, sampling, and the ideal paradigm for the study of causal relations. Chapter 4 classifies the types of questions that empirical research is asked to answer; this classification aids in deciding what types of research methods are appropriate for any given study. Chapter 5 explores the relationship between theory and empirical research. And Chapter 6 discusses the choice of appropriate empirical proxies (indicators) for theoretical variables.

Part Two gets down to the brass tacks of just how to conduct a piece of research. Chapter 7 provides a checklist of the basic steps one often takes when executing an empirical research project. If you are actually doing a
Other chapters in Part Two discuss crucial decisions in the research process—assessing the value of a prospective piece of research (Chapter 8), and whether to choose experimentation or the survey method (Chapters 10 and 13) or other methods (Chapter 14). Other chapters cover efficiency in sampling (Chapter 9), experimental design (Chapter 11), and the procedures of classification and measurement (Chapters 15 and 16). Part Two ends with the discussion of data handling and data adjustment in Chapter 17.

It was a toss-up whether to reverse Parts Two and Three. The reader may choose to read them in either order.

The chapters in Part Three consider the various types of obstacles that nature puts in the way of the fact seeker, and that prevent one from getting valid answers quickly and easily with common sense alone. Ways to surmount each obstacle are also discussed. These ways of surmounting obstacles to knowledge are the warp and woof of research method.

Part Four discusses what to do with your data once they are collected—how to analyze them and how to interpret them statistically (Chapters 25–30). Chapter 31 discusses how to decide on a sample size; it properly belongs in Part Two, but it had to follow after Chapters 25–30. Chapter 32 discusses the crucial concept of causality in social science.

3. The Place of Statistics in the Study of Research Methods

A working knowledge of the basic ideas of statistics and probability helps clarify one’s thinking and improves one’s capacity to deal with practical problems and to understand the world. And to be efficient a social scientist is almost sure to need knowledge of statistics and probability.

On the other hand, great research has been done by people with no formal knowledge of statistics. And a little study of statistics sometimes befuddles students into thinking that statistical principles are guides to research design and analysis. This mistaken belief only inhibits the exercise of sound research thinking. Kinsey put it this way:

However satisfactory the standard deviations may be, no statistical treatment can put validity into generalizations which are based on data that were not reasonably accurate and complete to begin with. It is unfortunate that academic departments so often offer courses on the statistical manipulation of human material to students who have little understanding of the problems involved in securing the original data. . . . When training in these things replaces or at least precedes some of the college courses on the mathematical treatment of data, we shall come nearer to having a science of human behavior. (Kinsey, et al., p. 35)

Throughout the book statistical ideas are submerged except as a handmaiden to research methods and research decisions. In addition, Chapters 27–31 offer a new approach to statistics that may have special interest for
students who are scared by statistics. This method substitutes “Monte Carlo” experiments for mathematical analysis. It also emphasizes the reasoning of statistics—the most important ideas in statistics, which are usually learned only informally. This method has now been shown experimentally to be unusually effective (Simon, Shevokas, and Atkinson).

4. Some General Remarks

A book about how to do research inevitably makes research sound difficult and treacherous. And this book gives special attention to the obstacles to knowledge that one faces in doing research, so as to help you recognize and overcome them.

But please don’t let these obstacles cause you to lose heart. Sound and valuable research can be done even by ordinary undergraduates—general research that is worth publishing, and applied research that is of value to organizations and decision makers. All research has flaws, but the flaws need not be so grave as to invalidate the research, even if it is course-work research conducted with a small sample, no money, and a limited budget of time.

Furthermore, research can be enormously exciting and great fun. To find out something about the world that no one has ever known before is a rare thrill. And it is a thrill that almost anyone can experience who will start with a sensible idea, work enthusiastically and hard, and proceed with caution.

Though research can be great fun, I do not like to think of it as a game. Rather, I prefer to remember that sound research can make a valuable social contribution, improving the lives of individuals and communities and enriching our culture. Thinking of research as a game can lead the researcher to focus only on the professional acceptance of one’s work and its influence on one’s career, rather than on the social and intellectual benefits of the research. Of course we all want to get ahead in the world. But if we want only to get ahead, if our eye is on only the main chance professionally, then we will all be losers in the long run.

This book is a textbook. Though some of the ideas in it are new, most are not. Like other textbooks it constitutes a sort of folk wisdom; the folk are the teachers, colleagues, and students who have discussed research with me. Some of this wisdom seems never to have been collected or transcribed from the oral tradition. (For example, the phrase *ceteris paribus* must have a very high spoken frequency among social scientists, yet it does not appear in the index of a single one of the most popular books about research methods in the social sciences.) To collect and discuss this wisdom is the aim of this book.
References


Klein, D. Short Notes. The Reporter of Direct Mail Advertising (February 1965), 4.525


Orcutt, Guy H. Actions, Consequences, and Causal Relations. Review of Economics and Statistics, XXXIV, No. 4 (November 1952), 305-313. (b)


Simon, Julian L. □□□Business Ethics, Business Ideology, and Business Behavior.□□□ Business and Society, 7 (Spring 1967), 18-32. (a)
Simon, Julian L. □□□Some Principles of Practical Welfare Economics.□□□ Management Science, 13 (June 1967), B621-629. (b)
Simon, Julian L. Unpublished ms. N.d.
Smith, Vernon L. □□□An Experimental Study of Competitive Market Behavior.□□□ Journal of Political Economy, LXX (April 1962), 111-137.534


Sudman, Seymour. The Uses of Telephone Directories for Survey Sampling. Journal of Marketing Research, 10 (May 1973), 204-207.


Wold, Herman O. A. Time as the Realm of Forecasting. Paper presented at Symposium on the Interdisciplinary Perspectives of Time, New York Academy of Sciences, January 20, 1966. (c)


Young, Pauline V. Scientific Social Surveys and Research: An Introduction to the Background, Content, Methods, and Analysis of Social Studies, 4th ed. New York: Prentice-Hall, 1966.


