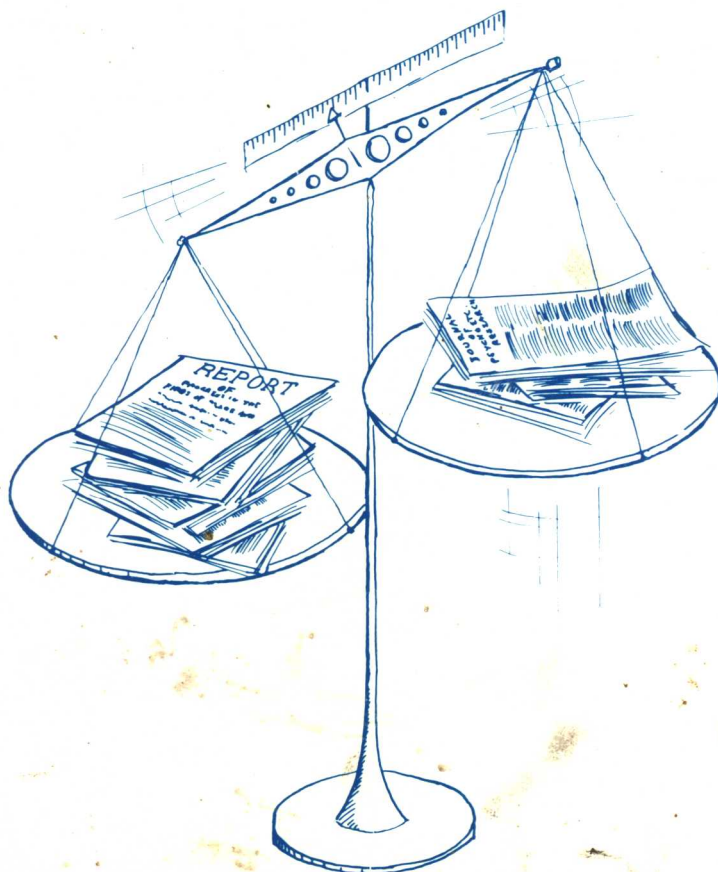

EVALUATING INFORMATION

**A Guide for Users
of Social Science Research**

Second Edition

**Jeffrey Katzer
Kenneth H. Cook
Wayne W. Crouch**



EVALUATING INFORMATION

**A Guide for Users
of Social Science Research**

Second Edition

Jeffrey Katzer

Syracuse University

Kenneth H. Cook

Data General World Headquarters

Wayne W. Crouch

Babson College

Random House



New York

Second Edition

9

Copyright © 1978, 1982 by Newbery Award Records, Inc

All rights reserved under International and Pan-American Copyright Conventions. No part of this book may be reproduced in any form or by any means, electronic or mechanical, including photocopying, without permission in writing from the publisher. All inquiries should be addressed to Newbery Award Records, Inc., 201 East 50th Street, New York, N.Y. 10022. Published in the United States by Newbery Award Records, Inc., a subsidiary of Random House, Inc., New York and simultaneously in Canada by Random House of Canada Limited, Toronto. Distributed by Random House, Inc

Manufactured in the United States of America

Library of Congress Cataloging in Publication Data

Katzer, Jeffrey.

Evaluating information.

Includes bibliographical references and index.

1. Social sciences—Research—Evaluation. 2. Social sciences—Methodology. I. Cook, Kenneth H. II. Crouch, Wayne W. III. Title.

H62.K378 1982 300'.72 81-3588

ISBN 0-394-34842-7

AACR2

PREFACE

This book was originally written to help offset an error we had been making as teachers. We had been giving our undergraduate students, our master's students (in a professional program), and our doctoral students (in a research program) the same type of introduction to social science research. All were taught from a researcher's point of view. Such a perspective is appropriate for those few students who will actually do research, but we had become convinced that it was not the best perspective for the many who will rarely, if ever, conduct a research investigation.

We are pleased that so many others found our approach from the user's point of view valuable. The first edition was adopted at over 353 colleges and universities in diverse fields of the social sciences. Thus, the major focus of the book has not changed in this second edition. We still have in mind undergraduate and graduate students in all of the social sciences and related professional areas, such as the following:

advertising	education	nursing
anthropology	geography	political science
broadcasting	journalism	psychology
communication	labor relations	public relations
counseling	library science	social work
criminal justice	management	sociology

What these students need is a research course with a consumer's point of view. In their coursework and in their professional careers these students will be called on to read and make use of social research findings, yet a researcher-oriented methods course of one or two semesters is not well-suited to their needs.

One approach is to develop a separate course for these students—a course in which the primary aim is to teach students how to understand and evaluate published research reports. This book can serve as the major textbook for such a course. However, we did not expect, and have not found, that many such courses exist.

A more viable approach is to add to existing courses some training and practice in research evaluation. Thus, we wrote this book primarily as a supplementary textbook. It can be used in courses on research methods and statistics and in courses that review research findings or make use of the original journal articles.

The wide applicability of the book is due in part to our approach. Our presentation is thematic—focusing on the broad concerns common to all forms of social research. Psychologists, for example, have traditionally relied on methods that are different from those of sociologists. Similarly, journalists have favored particular methods, while therapists have favored others. Common to all of these methods, however, are the same concerns, the same norms and values arising out of much of contemporary Western scientific practice and philosophy.

Most consumers of research will need to evaluate various types of research—experimental, survey, case study, and so on. Unlike researchers, consumers will not have much opportunity to become specialists in one or a few research methods. Consequently, they need a general orientation to what social research is about and a set of criteria that applies in most situations.

This book gives the reader the understanding needed to evaluate research reports. The bulk of the book is devoted to the principles underlying the criteria for evaluation. Each chapter leads to some specific questions that a reader must ask of a research report. At the end of the book we put these questions together in a step-by-step guide for evaluation and apply this guide in the evaluation of a journal article.

Because the book is aimed at nonresearchers, we were able to omit many of the technical terms, most of the fine distinctions, and all of the formulas that are so often important ingredients in a research methods textbook. Not only is this a nontechnical book but it may even be fun to read. We think that looking for clues as to whether information is factually accurate is challenging and rewarding—and if it isn't a pleasure it can at least be palatable, understandable, and useful.

The major difference between the first and second editions is the glossary; over 400 terms related to social research are defined and their relationships to other terms are indicated. We hope that, with the glossary, the book becomes more valuable as a continuing reference for, as well as an introduction to, the evaluation of social science research. We expect to see students using it as a handbook: they can look to the text for a refresher on basic concepts, to the step-by-step guidelines when they are faced with the need to read research reports, and to the glossary for a quick definition of an unfamiliar term.

Other changes in this edition include: (1) a clearer explication of the error model and the concept of noise; (2) substitution of the term *factually accurate* for *trustworthy* throughout the text; and (3) additions to Chapters 10, 12, and 14. Also, in the first edition arguments supporting generality were presented in a table in Chapter 13, but these same types of arguments can be

used to support a claim for factually accurate information. That table is now in Chapter 2; the arguments deal with both generality and factual accuracy.

Since the first edition was published many friends and colleagues have made suggestions, supplied citations, pointed out errors, and in many other ways have been very helpful. While we did not always implement their suggestions, we always considered them. The book is clearly better than it would have been if we had not received their input.

To some we must apologize. We have not added the many technical topics that were suggested. Unfortunately, there were almost as many technical topics to consider for addition as there were people suggesting them—each reviewer had a pet topic. Other reviewers suggested that we resist the temptation to make the text longer; we listened to them. We could not possibly add all of the technical topics suggested and still maintain the flavor of the first edition.

To the many who have influenced us (or tried to) we say thank you:

Rosemarie Bertran	Brenda Dervin	Craig B. Little
Burton Blatt	Donald P. Ely	Donald MacDonald
Harold Borko	Richard V. Farace	Gerald R. Miller
Richard Clark	Herbert Goldhor	David Ralph
Pauline Cochrane	Arnold Goldstein	John Simpkins
Michael Cody	Randy Harrison	James Sodt
William D. Coplin	Allan F. Hershfield	Albert Talbott
Cathy Covert	Dean E. Hewes	James A. Taylor
David Davidson	David R. Krathwohl	Robert S. Taylor
Clive Davis	Nan Lin	Douglas Zweizig

We are especially grateful to Walter Carroll, University of Bridgeport, who was kind enough to review the entire manuscript for the second edition. His perceptive comments encouraged us to modify what we had thought was a final draft and improve on it. Bill Frakes identified the terms for, and wrote the entire first draft of, the glossary. A special thank you for his technically accurate and monumental effort. Professor Allan Mazur helped us by reviewing the glossary at a time when we were too close to it to see it critically. We also want to thank Robert S. Sobel and Nancy Lillith for allowing us to reprint and evaluate one of their research studies.

Syracuse, New York
Medfield, Massachusetts
March 1981

J. K., W. W. C.
K. H. C.

CONTENTS

SECTION 1

INTRODUCTION:

How does anybody know anything?

- 1 Purpose: What we are trying to do 2
- 2 Assumptions: What we believe about how we know 10

SECTION 2

DISCOVERY AND COMMUNICATION OF RESEARCH FINDINGS:

Where do errors come from?

- 3 Observation: Seeing is not believing 22
- 4 Communication: Writing adds other problems 33
- 5 Interpretation: . . . And then you read it 43

SECTION 3

THE NATURE OF ERROR: What kinds are there?

- 6 Bias: A systematic error 50
- 7 Noise: the other type of error 59

SECTION 4

FACTUALLY ACCURATE INFORMATION: Can you believe it?

- 8 Subject matter: What is being studied? 68
- 9 Measurement: How does it size up? 86
- 10 Description: Are the results summarized fairly? 100
- 11 Control: Rival explanations . . . Is something else at work? 110
- 12 Inference: A real event . . . Or was it just luck? 126

SECTION 5

USEFUL INFORMATION:

Should you apply it?

13 Generality: Do the results apply to you? 144

14 Being practical: What else matters? 156

APPLICATIONS

How to do evaluations

A A step-by-step guide for evaluation 167

B Questions to ask 170

C A sample evaluation 175

- "Determinants of Nonstationary Personal Space Invasion,"
an article from the *Journal of Social Psychology* 176
 - Evaluation 183
-

A FINAL WORD 193

GLOSSARY 195

INDEX 233

INTRODUCTION

How does anybody know anything?

The social sciences have made great strides in advancing our understanding of people and social institutions. Yet with these advances also came some false starts, anomalous results, bad luck, and just plain mistakes. Unfortunately, there is no guaranteed way to prevent these errors from reaching print—for it is only with the certainty of hindsight that one can identify those studies that did not contribute to (or that even hindered) progress.

But how about the thousands of studies published each year that are too recent to be tested by time—how can they be evaluated? The approach we advocate in making this assessment is captured in what we call an “error model” of research. That model can not only be used by researchers to conduct their investigations, but it can be used equally well by readers of research to evaluate the published findings. We introduce this model in the first two chapters. The subsequent organization of the book as well as the particular criteria suggested for evaluation are based on this error model.

Thus, this first section of the book, though introductory, is essential. For unless you appreciate the need for evaluating the published research in the social sciences you will not have much motivation to read the remaining sections. And unless you accept our beliefs about how a researcher learns about the world, you will probably not accept our advice about how to evaluate what a researcher claims to have learned.

Purpose: What we are trying to do

Much information reported by scientists, published in reputable journals, and used by students, practicing professionals, and the general public is misleading. Some of it is just plain wrong. The purpose of this book is to help you detect such misinformation. It will help you become a more critical consumer of research published in the social sciences and in the related professional areas.

Students need to know how to evaluate research. More and more undergraduate courses include original research reports and journal articles as required readings. Course projects and term papers also often make use of published studies. Yet rarely are students taught how to evaluate what they need to read. Research courses exist, but they typically are introductions to how to *do* research, not how to independently *evaluate* it.

Professionals also need to know how to evaluate research. Included in this category are graduate students in those professional schools that make use of the theories, principles, and research findings of the social sciences. If you are, or are preparing to be, a social worker, educator, journalist, librarian, or manager you will not be conducting research very often during your career. Training in a profession requires time and coursework devoted to job preparation and practices; relatively little, if any, coursework is in research methods. But as a practicing professional you will be expected to keep up to date, to continually improve your skills and the knowledge by which you make decisions. Most new information will come from the ideas, suggestions, and research findings presented at professional meetings or published in professional journals. You must evaluate this new information and decide whether or not to act on the basis of it.

People in general need to evaluate information. Television bombards us with advertising claims. Popular magazines tout new diets and give tests to determine if you are really happy. Politicians and administrators want our support so they can change things for the better. Many times these appeals seem to be supported by research. But are they? In order to make sensible decisions, viewers, readers, and constituents must be able to evaluate the information presented.

The dilemma is clear. Whether you are a student trying to learn, a pro-



professional trying to do a better job, or a citizen trying to act intelligently, you need to read the published literature. Yet there is little in your training that will help you evaluate that literature.

THE LITERATURE IS PROBABLY WORSE THAN YOU THINK

Why should you have to evaluate research reports? Scientists have good reputations and the editor must have reviewed what was printed. It was published, so why not assume that it probably will be correct and applicable to current problems? There are many reasons why each article must be carefully read and evaluated. The various chapters of this book explore those reasons. But for the moment let us just consider the general accuracy or truthfulness of the published literature.

With a little thought it should become obvious that mistakes do happen, even to researchers, and that not all researchers or editors are equally knowledgeable or competent. Some errors will occur in published reports. Many people, however, would underestimate the seriousness of the errors, assuming that small mistakes in arithmetic or use of bibliographic citations or scientific equipment may be the problem: This book was written not because of such minor errors but because of the prevalence of errors that invalidate the major findings or conclusions of many studies. We will be concerned not with small errors, but with those poor research practices that leave the conclusions unsupported by the evidence or the arguments of the researcher. Some examples may help.

1. In a study of the quality of educational research, expert researchers evaluated a random sample of eighty-one articles published in thirty-

one different educational journals. Acting as if they were editors, they judged that only 7 percent of the articles were worthy of publication as written, 41 percent needed revision if they were to be published, and the remaining 52 percent of the published articles should not have been published at all.

2. In a carefully reasoned examination of the work of many contemporary and famous historians, one critic identified no fewer than one hundred different fallacies common and sometimes prevalent in their writings. These fallacies were often of such a nature as to seriously weaken the conclusions drawn.
3. In one small-scale study an attempt was made to obtain the actual data upon which some published articles were based. The original data for seven articles published in psychology journals were obtained. These data were re-analyzed and compared with the findings presented in the published articles. In three of the seven articles the second analysis found major mistakes that were not apparent in the published version. The mistakes, if known, would have altered the conclusions originally presented.
4. In a book devoted to examining errors in various economic fields the author found large and apparently unacknowledged errors in foreign trade statistics, prices, mining and agricultural production figures, employment statistics, national income statistics, and economic growth rate indexes. Use of these figures could be misleading to policy makers as well as to the casual reader.
5. Several other investigators found that practicing researchers in psychology, education, sociology, and communication are not fully knowledgeable about their craft. Many of these researchers believe things to be true about the process of research that are clearly not true.
6. "Dr. Richard W. Roberts, director of the National Bureau of Standards, estimates that half or more of the numerical data published by scientists in their journal articles are unusable because there is no evidence that the researcher accurately measured what he thought he was measuring or no evidence that possible sources of error were eliminated or accounted for."

Other examples can be given, but we think the point is clear. In all of the social sciences the literature contains more errors and erroneous conclusions than the average nonresearcher is usually led to believe. If the journals in the scientific fields where the better-trained researchers publish their findings are not 100 percent accurate, what is the situation in the more applied journals and in the magazines that popularize social research? An insidious possibility is that poorly conducted research will be published in these journals precisely because they do not meet the standards of the other ones. Furthermore, readers of popularized or applied social science may find these articles more believable because they are more understandable or because the article looks "scientific" and "statistical." In the 1940s, Ralph

A. Beals characterized the literature in library science as “glad tidings and testimony, with precious little research.” The increase in research-based literature has grown since that time in all professions. It does not always, however, offer anything more than glad tidings and testimony.

We do not mean through these examples and comments to condemn the various fields and professions cited. All journals in all fields publish articles that contain errors or are at some later time shown to be in error. We cite the fields that we do because the number and seriousness of errors have been studied. When we consider the findings of these studies along with the inadequate training in evaluating research that many nonresearchers in these fields receive, we believe there is ample reason to broadcast a warning: All readers in their own best interests need to be able to identify poorly conducted research. And, unfortunately, they cannot avoid that responsibility by relying on the prestige of the journal or the status of the author—neither is a guarantee and in some areas it may not even be a good rule of thumb. Nor can readers assume that an editor will successfully weed out all major errors and poorly conducted studies. The explosion of information, the reward structure in higher education, the increasing number of new journals being published, the evaluation procedures used by editors, and a variety of other factors make it necessary that each reader judge each article on its own merits.

WHAT THIS BOOK CAN DO FOR YOU AND HOW WE DO IT

The purpose of this book is to help you become a more critical consumer of research. We will focus on the published literature in the social and behavioral sciences and related professional areas, although evaluation of any purported factual information is subject to the same considerations. For instance, friends and colleagues cite facts in conversations, advertisers use the results of studies to promote their products, and textbooks and newspaper articles provide secondhand reports of the results of research. What you learn from this book can be applied to those situations as well.

This book is for readers of research, not researchers. In its preparation we wanted to omit all technical aspects or procedural details that would be important to a researcher, but we didn't want to sacrifice the principles upon which these methods are based. Moreover, we wanted to describe these principles in a nontechnical way. This is a very fine line to follow and we may have occasionally missed the mark. Because we do omit some more technical concerns you will not be able to detect all errors in the literature. In our experience, however, poor research often clearly demonstrates itself through a variety of errors. And most major errors can be detected by readers who understand the material in the remaining chapters of this book.

Our view of research is based on an “error model” and our approach to evaluation reflects that point of view. We believe that the researcher's job is first to identify, and then to remove or reduce, sources of potential error so that findings can be trusted. Consequently, it is the job of an evaluator to

challenge a research report by searching for possible errors missed by the author. Since we will focus on various sources of error throughout the book and often use examples of poorly conducted research, you might get the idea that we do not believe in research. That would be a false impression. Not only do we do research ourselves, but we also put a lot of faith in research as a very good method for obtaining accurate information.

We do find, however, that poor research practices serve as more interesting and more useful examples than do good practices. To interpret our concern with error as a reflection of our view of the usefulness of research would be a most unfortunate case of guilt by association—when in fact we associate with error only to eliminate it.

Thus, the major focus of this book will be on how well error has been identified, removed, and controlled in a particular research effort. We will try to provide the basis for you to understand and apply criteria for evaluating information. However, we cannot provide all the criteria you will need. As we pointed out, this is a nontechnical (and, by the way, short) book, but a more fundamental reason is that journal articles can be evaluated on both external and internal criteria. *External criteria* are all those that are based on comparisons of what is in the article with what is already known about the topic. You would want, for instance, to compare new findings with everything else you know about the subject. Do they agree with the results of previous research? Do they correspond with, or at least not contradict, firsthand experiences you have had? External factors are important in assessing the usefulness and credibility of a journal article, but this book can be of little help. Your own personal experiences and training in the subject matter are your primary tools.

Internal criteria, on the other hand, are those that are relatively independent of the specific subject matter and are mainly concerned with the adequacy of the methods used in the study itself. Most of this book is devoted to an examination of internal criteria. We will focus on how the authors discovered, supported, or proved their findings, conclusions, and recommendations. We will point out specific clues to look for.

Notice that we used the word *clues*. Our students frequently want to know if an article is correct. Unfortunately, there is nothing an author can write in an article or say in an oral presentation that will prove something to be correct unequivocally. The best of methods, meticulously carried out, sometimes lead to false findings. And the worst methods will sometimes produce correct results. Thus, finding clues of bad methodology is only the start. You must go beyond the identification of errors and estimate the consequences of those errors on the findings, conclusions, and recommendations presented. You can never be absolutely sure of your evaluation, but a reasonable estimate is far better than blind acceptance or rejection.

We are all tempted to damn the perpetrators of bad methodology, but we must be careful not to fall victim to what David Hackett Fisher calls the "fallacist's fallacy," which consists of any of the following *false* statements.

1. *An argument which is structurally fallacious in some respect is therefore structurally false in all respects.*
2. *An argument which is structurally false in some respect, or even in every respect, is therefore substantively false in its conclusion.*
3. *The appearance of a fallacy in an argument is an external sign of its author's depravity.*
4. *Sound thinking is merely thinking which is not fallacious.*
5. *Fallacies exist independent of particular purpose and assumptions.*

This may not be the worst fallacy, but it is the most difficult for the new critic to avoid.

Thus, we contend that you need to evaluate research because much published work is in error. This book will give you an understanding of the concepts and principles that help competent researchers discover accurate information about the world. Most importantly, we will help you use this understanding to evaluate their work by identifying specific clues to look for. These clues are presented as questions at the end of each chapter.

I'M LEARNING TO BE A SOPHISTICATED CONSUMER OF RESEARCH.



Resist the temptation to read only these questions. This book is not mainly a step-by-step guide for evaluating research, or a collection of the essential questions to ask. It is a presentation of concepts, of a way of thinking, and of an attitude to take as you approach a published report.

Evaluators must know how to choose criteria appropriately and apply them knowledgeably. That is what this book is about.

QUESTIONS TO ASK

At the end of each chapter we will summarize by giving you specific questions to ask about research and your reactions to it. Here are the first ones.

1. Are you too optimistic about the accuracy of this research report?

Most nonresearchers are. There are many reasons why each article you use must be carefully read and evaluated. Studies have shown that much published research is not completely accurate.

2. What do you already know about this topic—firsthand knowledge as well as the results of other research?

This book deals primarily with internal criteria—those that are relatively independent of the specific subject matter and are for the most part concerned with the methods used in a study. You must also be concerned about external criteria, which are all those based on comparisons of what is in the report with what is already known about the topic.

3. Are you too quick to condemn all aspects of the report because of some detected shortcomings?

Most studies, including those that were competently done, have some weaknesses. But the mere presence of a weakness does not invalidate an entire study. Judgment is needed; the baby shouldn't be thrown out with the bath water.

FOR FURTHER THOUGHT

1. When, if ever, can you rely on the reputation of a journal or author instead of evaluating the research itself? What is the trade-off in terms of your time and effort versus the danger of acting on the basis of erroneous information?
2. How would the literature of your field compare with the quality of published research cited in this chapter?
3. Were you surprised by the poor quality of much social science research literature? Why?
4. Interpret the following quotations in terms of the material in this chapter.
 - (a) "An educated person is one who has learned that information almost always turns out to be at best incomplete and very often false, misleading, fictitious, mendacious—just dead wrong." (R. Baker)
 - (b) "The positions for which social science students are likely to be preparing themselves—teaching, administration in government or business, community consultations, social work—increasingly call for the ability to evaluate and to use research results: to judge

whether a study has been carried out in such a way that one can have reasonable confidence in its findings and whether its findings are applicable to the specific situation at hand." (C. Selltiz, L.S. Wrightsman, and S. Cook)

FOR FURTHER READING

Our six examples of the inadequacy of social science research came from the following sources; they are listed in the same order here as used in the text. You might wish to evaluate those evaluations.

1. E. Wandt (Ed.). *A Cross-Section of Educational Research*. New York: McKay, 1965, pp. 1-7.
2. D. H. Fischer. *Historians' Fallacies: Toward a Logic of Historical Thought*. New York: Harper & Row, 1970, p. 305.
3. L. Wolins. "Responsibility for Raw Data." *American Psychologist* 17 (1962): 657-658.
4. O. Morgenstern. *On the Accuracy of Economic Statistics*, 2nd ed. Princeton, N.J.: Princeton University Press, 1963.
5. A. Tversky and D. Kahneman. "Belief in the Law of Small Numbers." *Psychological Bulletin* 76 (1971): 105-110.
 J. K. Brewer. "A Note on the Power of Statistical Tests in the American Educational Research Journal." *American Educational Research Journal* 9 (1972): 391-401.
 T. J. Duggan and C. W. Dean. "Common Misinterpretations of Significance Levels in Sociological Journals." *The American Sociologist* 3 (1968): 45-46.
 J. Katzer and J. Sodt. "An Analysis of the Use of Statistical Testing in Communication Research." *The Journal of Communication* 23 (1973): 251-265.
6. B. Rensberger. "Fraud in Research is a Rising Problem in Science." *New York Times*, January 23, 1977, p. 44.

Other fascinating examples of misinformation can be found in

T. Burnham. *Dictionary of Misinformation*. New York: Crowell, 1975; and *More Misinformation*. New York: McKay, 1980.