



Basic Content Analysis  
SECOND EDITION

---

**title:** Basic Content Analysis Sage University Papers Series. Quantitative Applications in the Social Sciences ; No. 07-049  
**author:** Weber, Robert Philip.  
**publisher:** Sage Publications, Inc.  
**isbn10 | asin:** 0803938632  
**print isbn13:** 9780803938632  
**ebook isbn13:** 9780585180847  
**language:** English  
**subject:** Social sciences--Methodology, Content analysis (Communication)  
**publication date:** 1990  
**lcc:** H61.W422 1990eb  
**ddc:** 300/.14  
**subject:** Social sciences--Methodology, Content analysis (Communication)

[cover](#)

[next page >](#)

SAGE UNIVERSITY PAPERS

Series: Quantitative Applications in the Social Sciences

Series Editor: Michael S. Lewis-Beck, *University of Iowa*

Editorial Consultants

Richard A. Berk, *Sociology, University of California, Los Angeles*

William D. Berry, *Political Science, Florida State University*

Kenneth A. Bollen, *Sociology, University of North Carolina, Chapel Hill*

Linda B. Bourque, *Public Health, University of California, Los Angeles*

Jacques A. Hagenaars, *Social Sciences, Tilburg University*

Sally Jackson, *Communications, University of Arizona*

Richard M. Jaeger, *Education, University of North Carolina, Greensboro*

Gary King, *Department of Government, Harvard University*

Roger E. Kirk, *Psychology, Baylor University*

Helena Chmura Kraemer, *Psychiatry and Behavioral Sciences, Stanford University*

Peter Marsden, *Sociology, Harvard University*

Helmut Norpoth, *Political Science, SUNY, Stony Brook*

Frank L. Schmidt, *Management and Organization, University of Iowa*

Herbert Weisberg, *Political Science, The Ohio State University*

Publisher

Sara Miller McCune, Sage Publications, Inc.

INSTRUCTIONS TO POTENTIAL CONTRIBUTORS

For guidelines on submission of a monograph proposal to this series, please write

Michael S. Lewis-Beck, Editor

Sage QASS Series

Department of Political Science

University of Iowa

Iowa City, IA 52242

[< previous page](#)

cover-0

[next page >](#)

Series/Number 07-049

Basic Content Analysis

SECOND EDITION

Robert Philip Weber

Harvard University



SAGE PUBLICATIONS

*The International Professional Publishers*  
Newbury Park London New Delhi

[< previous page](#)

page\_1

[next page >](#)

Copyright © 1990 by Sage Publications, Inc.

All rights reserved. No part of this book may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without permission in writing from the publisher.

*For information address:*



SAGE Publications, Inc.  
2455 Teller Road  
Newbury Park, California 91320  
E-mail: [order@sagepub.com](mailto:order@sagepub.com)

SAGE Publications Ltd.  
6 Bonhill Street  
London EC2A 4PU  
United Kingdom

SAGE Publications India Pvt. Ltd.  
M-32 Market  
Greater Kailash I  
New Delhi 110 048 India

Printed in the United States of America

International Standard Book Number 0-8039-3863-2

Library of Congress Catalog Card No. 90-061019

98 99 00 01 02 03 12 11 10 9 8 7

Sage Production Editor: Susan McElroy

When citing a University Paper, please use the proper form. Remember to cite the Sage University Paper series title and include the paper number. One of the following formats can be adapted (depending on the style manual used):

(1) HENKEL, RAMON E. (1976) *Tests of Significance*. Sage University Paper Series on Quantitative Applications in the Social Sciences, 07-004. Newbury Park, CA: Sage

OR

(2) Henkel, R. E. (1976) *Tests of significance* (Sage University Paper Series on Quantitative Applications in the Social Sciences, series no. 07-004). Newbury Park, CA: Sage.

Contents	
Series Editor's Introduction	5
Acknowledgements	7
1. Introduction	9
	10
Some Content-Analytic Studies	
	12
Issues in Content Analysis	
	13
Concluding Remarks	
	14
Suggestions for Further Reading	
2. Content Classification and Interpretation	15
	17
Reliability	
	18
Validity	
	21
Creating and Testing a Coding Scheme	
	24
Dictionaries and Computerized Text Classification	
	28
Reliability and Classification Schemes	
	32
Single Versus Multiple Classification	
	37
Assumed Versus Inferred Categories	
	38
Alternate Classification Schemes and Substantive Results	
	39
Units of Aggregation Problems	
	40
Concluding Remarks	
	40
Suggestions for Further Reading	
3. Techniques of Content Analysis	40
	42
Document Selection and Sampling	
	43
Text Encoding	
	44
Key-Word-In-Context Lists and Concordances	
	49

Word-Frequency Lists	53
Retrievals from Coded Text	56
Category Counts	58
Measurement Models	63
Accounting for Content 1: Characteristics of the Message Producers	64
Accounting for Content 2: Changes in the Socioeconomic System	68
Concluding Remarks	69
Suggestions for Further Reading	

4. Issues in Content Analysis	70
	70
Measurement	
	73
Indication	
	76
Representation	
	77
Interpretation	
	80
Concluding Remarks	
	80
Suggestions for Further Reading	
Appendix	
	80
Computer Software and Text Data Archives	
Notes	82
References	89
About the Author	96

## Series Editor's Introduction

*Content analysis* classifies textual material, reducing it to more relevant, manageable bits of data. Social scientists who must make sense of historical documents, newspaper stories, political speeches, open-ended interviews, diplomatic messages, psychological diaries, or official publications to name a few will find the technique indispensable. Take, as an example, political propaganda studies in the field of mass communications. Suppose Professor West wishes to examine the use of anti-Communism as a tool of political rhetoric in the speeches of American presidents. Therefore, he "content-analyzes" the annual State of the Union messages since 1948, counting the appearance of the word *Communism* in each.

This simple content analysis raises many questions. Is the word *Communism* a valid measure of the researcher's concept? Is it reliable? Should a broader code category (e.g., ANTI-COMMUNISM) be constructed, with any of a set of words (e.g., *Communism*, *Russia*, *Soviet threat*) counted? How does this quantitative indicator relate to foreign policy behavior? What about other themes in these messages?

Questions of this nature are treated in Dr. Weber's fine monograph, from the creation of a simple coding scheme to an elaborate computer-aided analysis of content. He makes his points with numerous well-chosen pieces of text: U.S. political party platforms, the 1886 address of the British King to parliament, speeches of the Kaiser, and Korean War editorials from American newspapers. Moreover, the utility of working by computer is spelled out. Once the text is computerized, say with an optical scanner, it is relatively easy to make a classification from more than one dictionary. Moreover, with computers, the coding rules are necessarily made explicit, allowing for perfect "intercoder reliability." Also, many of the computer options would be virtually impossible by hand, such as key-word-in-context (KWIC), which prints the context (the surrounding sentences) around each and every use of the selected keyword.

AUTHOR'S NOTE: *The first edition of this book was sponsored by the former series editors, Richard Niemi and John Sullivan.*

Conveniently, Dr. Weber includes an Appendix that considers different computer programs for the analysis of text. At the end of each chapter, he provides a good discussion of current literature. Hence, this second edition is fully up-to-date, in terms of substance and technology. Lest we forget, however, he sagely reminds us that content analysis is still "in part an art" (Chapter 3). Dr. Weber finishes with a sensitive discussion of the unresolved problems (in measurement, indication, representation, and interpretation) remaining for content analysis.

MICHAEL S. LEWIS-BECK  
SERIES EDITOR

[< previous page](#)

page\_6

[next page >](#)

## Acknowledgements

### Acknowledgements for Second Edition

Much of the material discussed in Chapter 4 will be covered in greater depth in a much longer piece coauthored with Zvi Namenwirth (Weber and Namenwirth, in press); Namenwirth's contribution, comments, and suggestions are greatly appreciated. Connie Zuell, Peter Philip Mohler, and Philip J. Stone provided continued encouragement and support, for which I am grateful. Much of the editing and rewriting was done in the *Patisserie Francaise* near Harvard Square. Many thanks to Elie and Salwa Matta and to Alexis (Gigi) Andrews for good food and hospitality.

### Acknowledgements for First Edition

This research was supported in part by ZUMA, the Center for Surveys, Methods, and Analysis, Mannheim, FRG. I remain indebted to the generosity of past and present Directors and staff, especially, Hans-Dieter Klingemann, Peter Philip Mohler, Karl Ulrich Mayer, Max Kaase, and Manfred Keuchler. Additional support was provided by Kurzweil Computer products, Inc., a subsidiary of Xerox Corp. Sue Williamson of Kurzweil played a key role, and her generosity and help is gratefully acknowledged. Randi Lynn Miller assisted with data entry at an earlier phase of the research. Thanks to Nancy Marshall, J. Zvi Namenwirth, Barbara Norman, Philip J. Stone, Kathrine Tenerowicz, and two anonymous reviewers for detailed comments on earlier drafts. Thanks also to the Harvard University Computing Center for computer resources.

Thanks to Zvi Namenwirth for permission to quote extensively from Namenwirth and Bibbee (1975); to Dexter Dunphy for permission to reproduce a figure from Dunphy, Bullard, and Crossing (1974); to Human Sciences Press for permission to adapt or reprint some material from Weber (1984b); to Elsevier Publishing Company for permission to adapt or reprint some material from Weber (1983); and to the Austrian Academy of Sciences for permission to adapt or reprint some material that appeared in Weber (1984a).

intentionally left blank

1.

## Introduction

Content analysis is a research method that uses a set of procedures to make valid inferences from text. 1 These inferences are about the sender(s) of the message, the message itself, or the audience of the message. The rules of this inferential process vary with the theoretical and substantive interests of the investigator, and are discussed in later chapters.

Content analysis can be used for many purposes. The following list points out a few notable examples (adapted from Berelson, 1952):

- disclose international differences in communication content;
- compare media or "levels" of communication;
- audit communication content against objectives;
- code open-ended questions in surveys;
- identify the intentions and other characteristics of the communicator;
- determine the psychological state of persons or groups;
- detect the existence of propaganda;
- describe attitudinal and behavioral responses to communications;
- reflect cultural patterns of groups, institutions, or societies;
- reveal the focus of individual, group, institutional, or societal attention; and
- describe trends in communication content.

The numerous examples presented throughout this monograph mainly show the last three uses of content analysis.

This monograph is an introduction to content analysis methods from a social science perspective.<sup>2</sup> The material covered here will be useful to students and researchers who wish to analyze text. The following