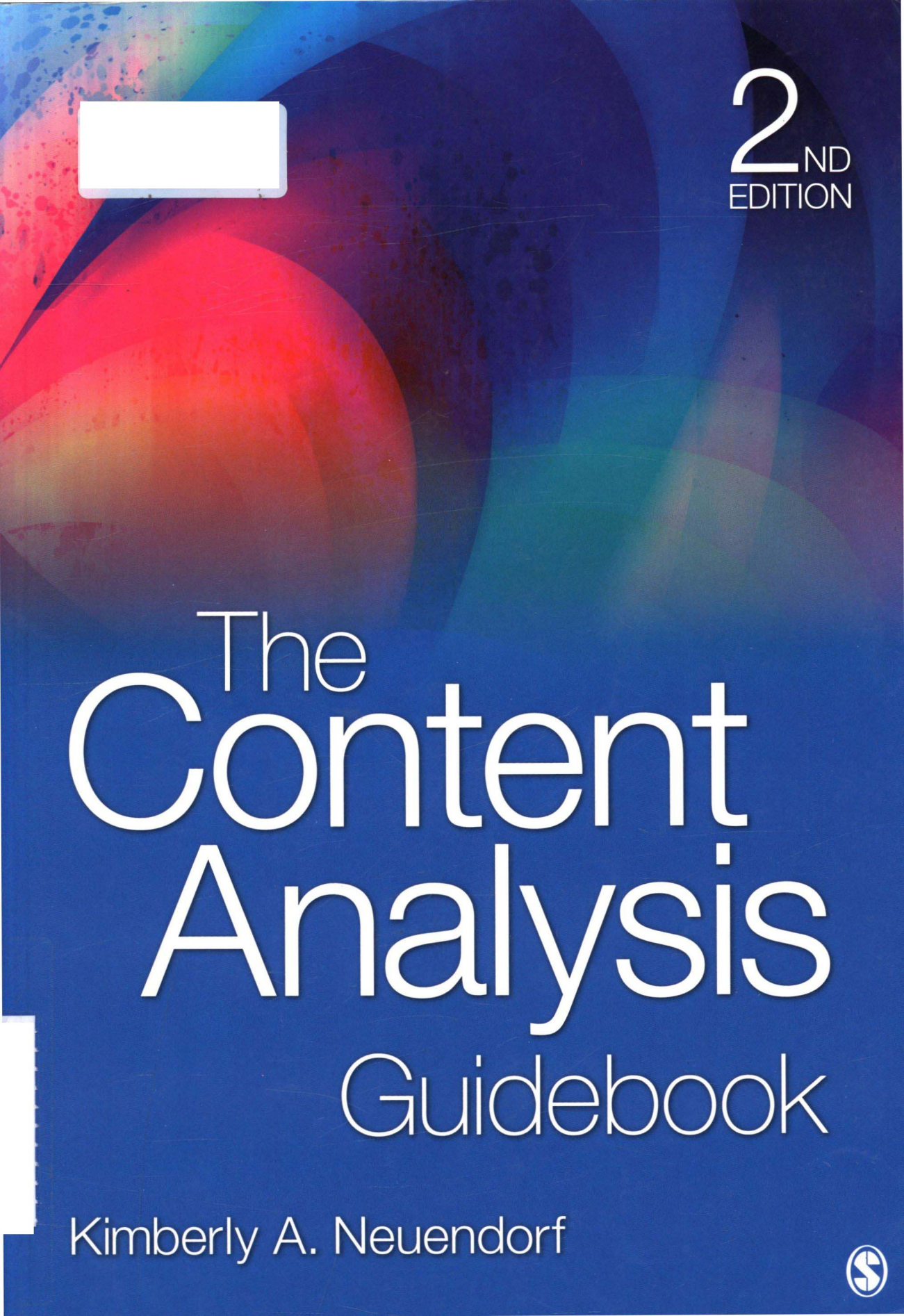




2ND
EDITION

The Content Analysis Guidebook

Kimberly A. Neuendorf



The Content Analysis Guidebook ²_{ND} EDITION

Content analysis is one of the most important but complex research methodologies in the social sciences. In this thoroughly updated second edition of **The Content Analysis Guidebook**, author Kimberly A. Neuendorf draws on examples from across numerous disciplines to clarify the complicated aspects of content analysis through step-by-step instruction and practical advice. Throughout the book, the author also describes a wide range of innovative content analysis projects from both academia and commercial research that provide readers with a deeper understanding of the research process and its many real-world applications.

New to the Second Edition

- **A new chapter on Content Analysis in the Interactive Media Age** (Chapter 7) shows readers how to create, acquire, archive, and code interactive media content.
- The **integrative model of content analysis**, which explains how content analysis may be linked with source and/or receiver characteristics, has been revised to clarify a difference between data links and logical links among source-message-receiver components.
- **New examples and updated references throughout** keep readers up-to-date with the latest scholarship in content analysis and its application to everyday life.
- **A new section focused specifically on validity** gives readers a deeper understanding of measurement and illustrates how the standards of validity interrelate.
- **A new resource section devoted to computer-aided text analysis (CATA) programs** such as Yoshikoder introduces readers to a growing set of options for automated analyses.

Cover image: thinkstock.com



Neuendorf The Content Analysis Guidebook

2ND
EDITION



The Content Analysis Guidebook

Second Edition

Kimberly A. Neuendorf
Cleveland State University



Los Angeles | London | New Delhi
Singapore | Washington DC | Melbourne



FOR INFORMATION:

SAGE Publications, Inc.
2455 Teller Road
Thousand Oaks, California 91320
E-mail: order@sagepub.com

SAGE Publications Ltd.
1 Oliver's Yard
55 City Road
London, EC1Y 1SP
United Kingdom

SAGE Publications India Pvt. Ltd.
B 1/1 Mohan Cooperative Industrial Area
Mathura Road, New Delhi 110 044
India

SAGE Publications Asia-Pacific Pte. Ltd.
3 Church Street
#10-04 Samsung Hub
Singapore 049483

Acquisitions Editor: Karen Omer
Editorial Assistant: Sarah Dillard
Production Editor: Libby Larson
Copy Editor: Amy Harris
Typesetter: C&M Digitals (P) Ltd.
Proofreader: Jennifer Grubba
Indexer: Jeanne R. Busemeyer
Cover Designer: Candice Harman
Marketing Manager: Amy Lammers

Copyright © 2017 by SAGE Publications, Inc.

All rights reserved. No part of this book may be reproduced or used 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.

Printed in the United States of America

Library of Congress Cataloging-in-Publication Data

Names: Neuendorf, Kimberly A., author.

Title: The content analysis guidebook / Kimberly A. Neuendorf, Cleveland State University, USA.

Description: Los Angeles : SAGE, [2017] | Earlier edition: 2002. | Second edition bibliographical references and index.

Identifiers: LCCN 2015044657 |
ISBN 9781412979474 (pbk. : alk. paper)

Subjects: LCSH: Sociology—Research—Methodology. |
Content analysis (Communication)

Classification: LCC HM529 .N47 2017 | DDC 301.01—dc23
LC record available at <http://lccn.loc.gov/2015044657>

This book is printed on acid-free paper.

16 17 18 19 20 10 9 8 7 6 5 4 3 2 1

The Content Analysis Guidebook

Second Edition

*To my family—Bob, Dorian, and Quinn, all of whom
contributed to the 2nd edition in their own way*

*In memoriam—In memory of my colleague and friend Paul D. Skalski, PhD, whose
contributions were many and whose spirit will never fade. His substantial abilities,
enthusiasm, and support were essential to both editions of this book.*

Preface

Content analysis has a history of more than 60 years of use in communication, journalism, sociology, psychology, business, and other disciplines. Its methods stem primarily from work in the social and behavioral sciences, but its application has reached such distant areas as law and health care. I've been involved with studies using the various methods of content analysis since the 1970s. Over that time, many things have changed, and others have remained amazingly constant. We now use computers to organize and analyze messages and to conduct statistical analyses with great speed. Yet studies are sometimes still conducted with little attention to theory or rigorous methods. There seems to be a wrong-headed assumption that anyone can do a content analysis with no training and little preparation. There is also contestation as to what types of message studies may be termed a *content analysis*. There remains a need for a clear and accessible text that defines the rules of the game and lays out the assumptions of this misunderstood quantitative research technique.

Forged through my own experiences as a coder, principal investigator, or adviser for at least 150 content analyses, I have developed a clear view of what content analysis can be when practiced with a "high bar." In my work, I have maintained a commitment to the centrality of content analysis to communication research and devoted my efforts to elevating the standards for content analysis. This book was written with two somewhat contradictory goals—to combine a strong scientific approach and high methodological standards with a practical approach that both academics and industry professionals will find useful. Resource 1 provides the reader with a guide to the comparison of text analysis computer programs. Other support materials can be found at the book's web site, *The Content Analysis Guidebook Online*—for instance, sample codebooks, coding forms, dictionaries, bibliographies, and more information on archives and computer text programs (see Resource 2).

This book is designed for upper-level undergraduates and graduate students studying communication, sociology, psychology, and other social sciences. It should also be useful to academics and practitioners in such related areas as marketing, advertising, journalism, film, literature, public relations, and other business-related fields, as well as all other areas that are concerned with the generation, flow, and impact of *messages*.

Acknowledgments_____

This book is the culmination of more than 35 years of research and teaching involving the method of content analysis. I owe a debt of gratitude to so many people whom I have encountered along the way.

First, special thanks go to Bradley S. Greenberg, the late Charles K. Atkin, and M. Mark Miller for granting me the initial opportunities at Michigan State University that have led to my expertise in the methodology of content analysis.

I'd like to acknowledge the influence of my long-time Cleveland State University colleagues Leo Jeffres, David Atkin, Sue Hill, and the late Sid Kraus, whose help in producing the first edition of this book was vital. The help of other CSU colleagues in providing counsel and support in the process of refining the second edition has been invaluable—thanks to Jill Rudd, Anup Kumar, Guowei Jian, Rick Pitchford, and the late Paul D. Skalski.

The professional staff at Sage Publications has been of great help. Particular thanks go to Matthew Byrnie (and his unending patience during difficult times) and copy editor Amy Harris.

Academic colleagues at other institutions have provided opportunities for expanding my vision of content analysis and have been generous with their input for the second edition. They include David Voelker, Tom Johnson, Julie Cajigas, Pam Shoemaker, Ben Bates, Ed Fink, Vish Vishwanath, and Limor Shifman. Of particular note are the important contributions of statistician James Henderson, PhD, to the substance of Chapter 6; I am forever grateful to him for his generosity. And contributions from members of the content analysis community have been enlightening; thanks go to James Pennebaker, Cindy Chung, Mike Thelwall, Joe Woelfel, Michael Young, Will Lowe, Cheryl Schonhardt-Bailey, David Fan, Sharon Smith, Cherie Chauvin, Kristen Harrison, George Barnett, Amy Jordan, Dale Kunkel, Jennifer Manganello, Matthew Lombard, Julie Andsager, and the late Phil Stone.

Special thanks go to students of mine who have been particularly instrumental in the expansion of my content analysis expertise into areas where I might never have ventured otherwise. Thank you to Jeffery “Phoenix” Allen, Dr. Jordan Tobin, Devin Kelly, Matt Egizii, Mu Wu, Ev Penev, Carolyn Kane, Serena Hendricks, Ben Gifford, Jeff Spicer, Rachel Campbell, Elizabeth

Langdon, Olga Prominski, Alicia Romano, Amy Brown, Pete Lindmark, Kim Cooper, Mary Casey, David Kasdan, Bettina Lunk Weber, Amy Dalessandro, Tom Gore, Patricie Janstova, Sharon Snyder-Suhy, Elizabeth Kenney, Ryan Weyls, Patinuch Wongthongsri, John Naccarato, Jeremy Kolt, Amy Capwell Burns, Barb Brayack, Jean Michelson, Ann Marie Smith, and James Allen Ealy. And sincere thanks to Dorian Neuendorf for her invaluable, incisive assistance in editing.

Contents

LIST OF BOXES	xi
LIST OF TABLES AND FIGURES	xiii
PREFACE	xv
ACKNOWLEDGMENTS	xvi
 Chapter 1: Defining Content Analysis	 1
An Introduction	1
The Growing Popularity of Content Analysis	2
The Myths of Content Analysis	5
Myth 1: Content Analysis Is Limited to Simple Analyses	5
Myth 2: Anyone Can Do Content Analysis; It Doesn't Take Any Special Preparation	7
Myth 3: The Term <i>Content Analysis</i> Applies to <i>All</i> Examinations of Messages	9
Myth 4: Content Analysis Is for Academic Use Only	15
A Six-Part Definition of Content Analysis	16
1. Content Analysis as Following the Standards of the Scientific Method	18
2. The Message as the Unit of Analysis, the Unit of Data Collection, or Both	20
3. Content Analysis as Quantitative	21
4. Content Analysis as Summarizing	23
5. Content Analysis as Applicable to <i>All</i> Contexts	24
6. All Message Characteristics Are Available to Content Analyze	31
 Chapter 2: An Integrative Approach to Content Analysis	 36
The Language of the Scientific Method	36
Human Coding Versus Computer	
Coding: A Brief Overview	39

How Content Analysis Is Done: A Flowchart for the Typical Process of Content-Analytic Research	39
Approaches to Content Analysis	42
Descriptive Content Analysis	43
Inferential Content Analysis	44
Psychometric Content Analysis	45
Predictive Content Analysis	46
Toward an Integrative Model of Content Analysis	50
Using the Integrative Model of Content Analysis	52
Integrative Data Linking	53
Integrative Logical Linking	54
The $S \rightarrow M$ Data Link:	
Linking Source and Message Data	54
The $S \rightarrow M$ Logical Link:	
Linking Source and Message Findings	59
The $M \rightarrow R$ Data Link:	
Linking Message and Receiver Data	60
The $M \rightarrow R$ Logical Link:	
Linking Message and Receiver Findings	63
The Growing Use of Integrative Studies	65
Developing New Data Linkages	67
Chapter 3: Message Units and Sampling	70
Units	70
Unitizing a Continuous Stream of Information	72
Defining the Population	74
Cross-Sectional Versus Longitudinal Populations	77
Availability-Based Versus	
Exposure-Based Populations	78
Archives	80
The Evaluation of Archives	83
Sampling	84
Probability Sampling	84
Nonprobability Sampling	89
Sample Size	90
Full Reportage	91
Chapter 4: Variables and Predictions	96
Identifying Critical Variables	97
A Consideration of Universal Variables	98
Using Theory and Past Research	
for Variable Collection	101
A Grounded or Emergent Process	
of Variable Identification	108
Attempting to Find	
Medium-Specific Critical Variables	110

Hypotheses and Research Questions	113
Conceptual Definitions	114
Hypotheses	115
Research Questions	117
Full Reportage	119
Chapter 5: Measurement and Validity	121
Defining Measurement	121
Validity, Reliability, Accuracy, and Precision	122
Validity	122
Reliability	122
Accuracy	123
Precision	123
How the Standards Interrelate	123
Types of Validity Assessment	124
Internal Validity and External Validity	125
Face Validity	125
Criterion Validity	126
Content Validity	127
Construct Validity	127
Validation Procedures in Content Analysis	129
Validation in Human-Coded Analyses	129
Validation in CATA Analyses	129
Ethics in Content Analysis	130
Operationalization	131
Categories or Levels That Are Exhaustive	131
Categories or Levels That Are Mutually Exclusive	131
An Appropriate Level of Measurement	133
Expanding Terminology and Merging Traditions:	
Big Data, Text Mining, Natural Language Processing, Computational Linguistics, Sentiment Analysis/Opinion Mining, and CATA	143
CATA (Computer-Aided Text Analysis)	146
Dictionaries for Text Analysis	148
CATA Versus Human Coding	153
Selection of a Computer Text Content Analysis Program	154
Human Coding	156
Codebooks and Coding Forms	156
Coder Selection	157
Coder Training	158
The Processes	159
Medium Modality and Coding	159
Index Construction in Content Analysis	160
Full Reportage	161

Chapter 6: Reliability	165
Intercoder Reliability: An	
Introduction to Standards and Practices	166
Issues in the Assessment of Reliability	169
Agreement Versus Covariation	169
Reliability as a Function of Subsamples and Coders	169
Threats to Reliability	170
Reliability for Manifest Versus Latent Content	170
Pilot and Final Reliabilities	172
Intercoder Reliability Coefficients: Issues and Comparisons	173
Agreement	174
Agreement Controlling	
for the Impact of Chance Agreement	175
Covariation	178
Calculating Intercoder Reliability Coefficients	180
The Reliability Subsample	187
Subsample Size	187
Sampling Type	187
Assignment of Cases to Coders	188
Treatment of Variables That Do Not	
Achieve an Acceptable Level of Reliability	188
The Use of Multiple Coders	189
Intracoder Reliability—Assessing Stability Over Time	190
Unitizing Reliability	191
Calculating Reliability Coefficients:	
Programs and Calculators	191
Special Issues in Reliability Coefficient Decision-Making	191
Alternative Coefficients	191
Standard Errors and Confidence Intervals	194
Controlling for Covariates	194
Sequential Overlapping Reliability Coding	194
Bringing Order to Reliability: Building	
Models That Specify Sources of Variation	195
Full Reportage	196
Chapter 7: Content Analysis in the Interactive Media Age	201
<i>Paul D. Skalski, Kimberly A. Neuendorf, and Julie A. Cajigas</i>	
Considering Interactive Media	207
Content Creation in the Interactive Media Age	207
The Special Nature of	
Content in the Interactive Media Age	209
Practical Suggestions	211
Acquiring Content in the Interactive Media Age	212
Acquiring Video Game Content	214
Acquiring Web 2.0 Content	217
Acquiring Online News	223
Content Acquisition by Interactive Media	223

Archiving Content in the Interactive Media Age	224
Archiving Video Games	224
Archiving Web 2.0 (and Other Web) Sites	224
Traditional Media Archiving	
With Interactive Technology	225
Coding Content in the Interactive Media Age	226
Unitizing Interactive Content	226
Codebooks and Coding Forms	227
Coding Interactive Content	228
The Use of Technologies for Medium	
and Message Management, Message Preparation	229
Automated Technologies and	
Measurement on the Moving Image	233
Almost Content Analysis	235
Conclusion	236
Chapter 8: Results and Reporting	243
Data Handling and Transformations	243
Hypothesis Testing	244
Hypotheses and Research Questions—A Reminder	244
Descriptive, Inferential,	
Parametric, and Nonparametric Statistics	244
Selecting the Appropriate Statistical Tests	246
Frequencies	249
Co-Occurrences and In-Context Occurrences	251
Timelines	253
Bivariate Relationships	254
Multivariate Relationships	259
Chapter 9: Contexts	273
Psychometric Applications of Content Analysis	274
Thematic Content Analysis	274
Clinical Applications	275
Open-Ended and Pictorial Responses	276
Linguistics and Semantic Networks	278
Stylometrics and Computer Literary Analysis	279
Interaction Analysis	280
Other Interpersonal Behaviors	282
Computer-Mediated Communication	283
Organizational Communication	283
Health Communication	285
Advertising	287
News	288
Political Communication	289
Violence in the Media	290
Gender Roles in the Media	292
Sex and Sexuality in the Media	294

Minority Portrayals	294
Movies	295
Music	297
Sentiment Analysis	297
Academic Literatures	298
Commercial, Governmental, and Other	
Client-Based Applications of Content Analysis	299
Legal Applications of Content Analysis	301
Future Directions	301
Resource 1: CATA—Computer-Aided Text Analysis Options	304
<i>Kimberly A. Neuendorf, Jeffery “Phoenix” Allen, Paul D. Skalski, and Julie A. Cajigas</i>	
Part I. Computer-Aided Text Analysis (CATA) Programs	307
CATPAC II	307
Concordance 3.3	308
Diction 7	308
General Inquirer	310
Hamlet II 3.0	310
LIWC2015	311
MCCALite	312
PCAD	312
Profiler Plus	313
SALT 2012	314
SentiStrength 2.2	315
TextAnalyst	316
Text Analytics for Surveys 4.0 (IBM SPSS)	316
TEXTPACK	317
TextQuest 4.2	318
T-LAB Pro 9.1.3	318
WordSmith 6.0	319
WordStat 7.0	320
Yoshikoder	320
Part II. Yoshikoder Example	321
Resource 2: The Content Analysis Guidebook Online (CAGO)	326
Content Analysis Resources	326
Message Archives and Corpora	326
Reliability	327
Human Coding Sample Materials	327
Introduction to Computer Content Analysis (CATA)	327
REFERENCES	328
AUTHOR INDEX	404
SUBJECT INDEX	422
ABOUT THE AUTHORS	438