

FIFTH EDITION

ANALYZING MEDIA MESSAGES

Using Quantitative Content Analysis in Research

Daniel Riffe, Stephen Lacy,
Brendan R. Watson, and Jennette Lovejoy



Analyzing Media Messages

The fifth edition of this comprehensive and engaging text guides readers through the essential tools and skills necessary to conduct quantitative content analysis research.

Readers will find a clear definition of quantitative content analysis and step-by-step instructions on designing a content analysis study, along with examples of content analysis studies and journal articles. This edition has been updated with the latest methods in sampling in the digital age, computerized content analysis, and the uses of social media in content analysis research. It maintains the concise, accessible approach of previous editions while including refreshed examples and discussions throughout.

This is an essential text for content analysis courses in communication and media studies programs at all levels, as well as a useful supplementary text in more general research methods courses.

Daniel Riffe is Professor Emeritus in the School of Journalism and Media at the University of North Carolina at Chapel Hill, USA, and former editor of *Journalism and Mass Communication Quarterly*. His research examines mass communication and environmental risk, political communication and public opinion, international news coverage, and research methodology. Before joining UNC-Chapel Hill, he was Presidential Research Scholar in the Social and Behavioral Sciences at Ohio University.

Stephen Lacy is Professor Emeritus at Michigan State University, USA, where he studied content analysis and media managerial economics for more than 30 years in the School of Journalism and Department of Communication. He has co-written or co-edited five other books and served as co-editor of the *Journal of Media Economics*.

Brendan R. Watson is an independent scholar. His research examines the role of public affairs news/information in helping communities cope with social upheaval due to the increasing urbanization, globalization, and pluralism of postindustrial society. He also studies research methodology. He has taught graduate seminars in content analysis at Michigan State University and the University of Minnesota, where he was previously on the faculty. He holds a Ph.D. in Mass Communication from the University of North Carolina at Chapel Hill. He currently works for Level, a performance marketing agency.

Jennette Lovejoy is Professor and Chair of the Department of Communication and Media at the University of Portland, USA. She teaches critical media, journalism, and research methods. Her interdisciplinary research spans media, health, and content analysis methods. She has edited multiple books and published in leading communication and medical journals.



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

Analyzing Media Messages

Using Quantitative Content
Analysis in Research

Fifth Edition

**Daniel Riffe, Stephen Lacy,
Brendan R. Watson, and
Jennette Lovejoy**

Designed cover image: BlackJack3D/© Getty Images

Fifth edition published 2024

by Routledge

605 Third Avenue, New York, NY 10158

and by Routledge

4 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

Routledge is an imprint of the Taylor & Francis Group, an informa business

© 2024 Daniel Riffe, Stephen Lacy, Brendan R. Watson, and
Jennette Lovejoy

The right of Daniel Riffe, Stephen Lacy, Brendan R. Watson, and
Jennette Lovejoy to be identified as authors of this work has been
asserted in accordance with sections 77 and 78 of the Copyright, Designs
and Patents Act 1988.

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

Trademark notice: Product or corporate names may be trademarks or
registered trademarks and are used only for identification and explanation
without intent to infringe.

First edition published by Routledge 1998

Fourth edition published by Routledge 2019

Library of Congress Cataloging-in-Publication Data

Names: Riffe, Daniel, author. | Lacy, Stephen, 1948– author. |

Watson, Brendan R., author. | Lovejoy, Jennette, author.

Title: Analyzing media messages : using quantitative content analysis
in research / Daniel Riffe, Stephen Lacy, Brendan R. Watson and
Jennette Lovejoy.

Description: Fifth edition. | New York, NY : Routledge, 2024. |

First edition published by Routledge 1998. Fourth edition published
by Routledge 2019. | Includes bibliographical references and
index. Identifiers: LCCN 2023028911 (print) | LCCN 2023028912
(ebook) | ISBN 9781032264691 (hardback) | ISBN 9781032264677
(paperback) | ISBN 9781003288428 (ebook)

Subjects: LCSH: Content analysis (Communication) | Mass media–
Research–Methodology. | Mass media–Statistical methods.

Classification: LCC P93 .R54 2024 (print) | LCC P93 (ebook) |

DDC 302.2301/4–dc23/eng/20231016

LC record available at <https://lcn.loc.gov/2023028911>

LC ebook record available at <https://lcn.loc.gov/2023028912>

ISBN: 978-1-032-26469-1 (hbk)

ISBN: 978-1-032-26467-7 (pbk)

ISBN: 978-1-003-28842-8 (ebk)

DOI: 10.4324/9781003288428

Typeset in Times New Roman

by Apex CoVantage, LLC

Daniel Riffe:

**For Florence, Ted, Eliza, Bridget, Brynne, Hank, Annie,
James, and Elizabeth**

Stephen Lacy:

For I. P. Byrom, N. P. Davis, and A. G. Smith

Brendan R. Watson:

For Joan and Maroun

Jennette Lovejoy:

For Georgia, Sawyer, and Roan



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

Contents

<i>Preface</i>	<i>viii</i>
1 Introduction	1
2 Defining Content Analysis as a Social Science Tool	21
3 Designing a Content Analysis	36
4 Computers and Content Analysis	57
5 Measurement	71
6 Sampling	92
7 Reliability	118
8 Validity	152
9 Data Analysis	168
<i>Appendix A: Sample Protocol</i>	<i>192</i>
<i>Appendix B: Reporting Standards for Content Analysis Articles</i>	<i>197</i>
<i>References</i>	<i>199</i>
<i>Index</i>	<i>220</i>

Preface

The purpose of the fifth edition of this book is to help facilitate the development of a science of communication, in particular as it relates to different forms of mediated communication. A communication science is at the heart of all our social sciences because communication increasingly defines what we do, how we do it, and even who we are individually, socially, and culturally.

In fact, never before in human history has mediated communication been so central, pervasive, and important to human civilization. A good communication science is necessary if humanity is to fully understand how communication affects us. Absent good understandings from such a communication science, we will always be at the mercy of unintended, unforeseen consequences.

But absolutely necessary to the development of a communication science is a means of logically assessing communication content. Broadly speaking, communication content varies based on a large set of factors that produce and deliver that communication. And, in turn, the variations in communication content affect a large set of individual, group, institutional, and cultural factors. In other words, understanding communication content is necessary and central to any communication science in which the goal is to predict, explain, and potentially control phenomena (Reynolds, 1971).

More specifically, we believe the systematic and logical assessment of communication content requires quantitative content analysis, the topic of this book. Only this information-gathering technique enables us to illuminate patterns in communication content reliably and validly. And only through the reliable and valid illumination of such patterns can we hope to illuminate content causes or predict content effects.

We bring to this effort our experiences conducting or supervising hundreds of quantitative content analyses in our careers as researchers, examining content ranging from White House coverage, to portrayal of women and minorities in advertising, to the sources given voice in local government news. The content analyses have included theses and dissertations, class projects, and funded studies, and have involved content from sources as varied as newspapers, broadcast media, Twitter, and websites. Some projects have been descriptive, whereas

others have tested hypotheses or sought answers to specific research questions. They have been framed in theory about processes that affect content and about the effects of content.

If conducting or supervising those studies has taught us anything, it is that some problems or issues are common to virtually all quantitative content analyses. Designing a study raises questions about accessing content, sampling, measurement, reliability, and data analysis—fundamental questions that arise whether the researcher is a student conducting her first content analysis or a veteran planning her 20th, whether the content being studied is words or images, and whether it comes from social networking sites or a legacy medium.

In preparing this book for the fifth edition, we re-engage these recurring questions. Our goal is to make content analysis accessible, not arcane, and to produce a comprehensive guide that is also comprehensible. We hope to accomplish the latter through clear, concrete language and by providing numerous examples—of recent and “classic” studies—to illustrate problems and solutions. We see the book as a primary text for courses in content analysis, a supplemental text for research methods courses, and a useful reference for fellow researchers in mass communication fields, political science, and other social and behavioral sciences.

This fifth edition varies from the previous four because a new coauthor, Jennette Lovejoy, has joined the team, while original author Frederick Fico has stepped down. In addition to Fico, we owe thanks to many for making this book possible: teachers who taught us content analysis (Donald L. Shaw, Eugene F. Shaw, Wayne Danielson, and James Tankard); colleagues who provided suggestions on improving the book; and our students who taught us the most about teaching content analysis. Jennette and Brendan learned content analysis by studying previous editions of this very book and doing content analysis with their mentors, with whom they are now co-authors.

Finally, our deepest appreciation goes to our families, who often wonder whether we do anything but content analysis.

Daniel Riffe
Stephen Lacy
Brendan R. Watson
Jennette Lovejoy



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

1 Introduction

Of all the social science methods (e.g., experiments, focus groups, surveys, etc.) available for use in researching the broad domain of communication, content analysis is intuitively the most central: “Because all human verbal and mediated exchanges involve messages (content), content analysis is particularly important for the study of communication” (Lacy, Watson, Riffe, & Lovejoy, 2015, p. 807).

One of the authors recalls early impressions of content analysis, based on published studies in graduate school half a century ago. Content analysts worked in libraries looking at printed (or microfilmed) newspaper pages to measure space given to a particular topic or person. While such studies, describing messages produced by professional communicators in covering particular topics, have merit and import for communication science, things have changed obviously since a half-century ago.

The “broad domain” for contemporary study requires reconceptualization of what is a *medium* of communication (e.g., ranging from legacy public news media in print, visual, or digital forms, to neighborhood, community, or corporate websites, to social media apps targeting broad or narrow audiences); who is a *communicator* (ranging from traditional trained professionals, to individuals using social media to influence followers or represent their real or “ideal” selves, to interest groups or communities that form via social media to address common interests or goals); and what constitutes a *message* (ranging from a Fortune 500 company’s corporate mission statement, to a newspaper editorial, to legacy media news posted on Facebook, to a candidate’s speech, a hate group’s online manifesto, an Instagram “selfie,” or a posted response to any of these).

Consider, then, the diversity of these quantitative content analyses.

With early COVID-19 infection rates three times greater among Black Americans than among Whites, and Black Americans twice as likely to die from the virus, Biswas, Sipes, and Brost (2021) compared Spring 2020 general media and Black media coverage of Black-related COVID-19 issues. Contrary to the authors’ hypothesis, Black media did not include more “social responsibility”

2 Introduction

frames (e.g., blaming systemic racism and inequality for unequal healthcare). However, a significantly greater percentage of Black media items (64%) included “action or solution frames” (e.g., safety guidance, testing options, etc.). When analyses focused specifically on Black-majority cities, general media were more likely to use social responsibility and consequence (e.g., health, political, economic) frames, while Black media were more likely to use the action/solution and individual responsibility (e.g., action by the President or a specific local official) frames.

After several popular rappers publicly disclosed battles with anxiety, depression, and other mental health conditions, scholars (Kresovich, Collins, Riffe, & Carpentier, 2021) explored mental health themes in lyrics of *Billboard* chart-topping rap songs, noting a significant two-decade increase. A plurality (28%) of songs referenced anxiety, 22% mentioned depression, and 8% alluded to suicide. “Contributing stressors” included the artists’ social environment and unhappy love life. Kresovich et al. pondered what effects these “increasingly prevalent messages may have in shaping mental health discourse and behavioral intentions” (p. 286).

Through the lens of “public diplomacy” (Golan, 2013)—in which “governments communicate and build relationships with foreign publics in pursuit of political objectives” (Fitzpatrick, Fullerton, & Kendrick, 2013, p. 1)—Sobel, Riffe, and Hester (2016) collected a year’s Twitter feeds from four embassies on the US State Department’s “watch list” (designated as “dangerous or unstable”: Afghanistan, Libya, Nigeria, and Syria) and four not on the list. A simple random sample of 25% of each embassy’s tweets was drawn, for a total of 2,625. Among “watch list” feeds, only the embassy in Syria commented on the ongoing civil war; Embassy Kabul was silent on the conflict in Afghanistan. Sobel et al. concluded there was little consistency among embassies in “formally furthering the State Department mission” (p. 102).

Video-game researchers (Lynch, Tompkins, van Driel, & Fritz, 2016) looked at female character “sexualization” across three decades—a period encompassing the 1996 *Tomb Raider* game that introduced Lara Croft, a character described as highly sexualized yet strong, bold, educated, and capable (p. 569). Sexualization increased from 1992 to 2006, but declined from 2007 to 2014. Lynch et al. reported a persistent relationship between sexualization and capability, a fact that may help “empower female gamers” (p. 578), though female characters were more often in secondary roles (p. 580).

Bastien (2018) compared newspaper coverage with verbatim transcripts from televised debates in five Canadian federal campaigns (1968–2008). Debate reports became increasingly “analytical and judgmental” and less “factual”: the presence of journalists’ opinions in paragraphs increased from 14% to 24% (p. 9). However, agendas of politicians and journalists were correlated: “the longer an issue is debated by the leaders, the more it is reported” (p. 1757).

Lee and Riffe (2017) explored how corporations and an industry monitoring group focus media attention on corporate social responsibility (CSR) activities (e.g., efforts to improve the environment, community, and employees). Data from 7,672 press releases from 223 US corporations, 1,064 *New York Times* and *Wall Street Journal* articles, and ratings of corporations by a CSR monitoring group showed stronger relationships between ratings and news coverage than between press releases and coverage. Companies may need to heed such monitoring groups and reconsider what they provide in press releases. Indeed, Ki and Hon (2006) earlier explored Fortune 500 companies' website promotion of CSR activities involving education, the community, and the environment, finding that few sites communicated effectively about CSR.

While many of these content analyses focused on the presence or *representation* of individuals, groups, and ideas *in* messages, other scholars have analyzed social media content (numbers of readers, sharing/linking, networking, etc.) to illustrate how communities of ideas are *created* by communicating via social media, and how “dominant media narratives” can be reshaped by that discourse. Indeed, research shows some editors and journalists attend to social media and alter content based on social media responses (Tandoc & Vos, 2016).

Qualitative research has examined “right-wing populism” (RWP) in social media platform content (e.g., Engesser, Ernst, Esser, & Büchel, 2017), while other researchers have explored the function of online RWP for followers. Heiss and Matthes (2020) captured 13,358 Facebook posts from political parties, candidates, and their followers in Austria and Germany, seeking anti-immigrant and anti-elite sentiments. They created “dictionaries” of anti-immigrant and anti-elite terms to use with a computerized coding process (see Chapter 4, this volume, for a discussion of automated approaches). Confirming that RWP party sites had more anti-immigrant and anti-elite references than non-RWP sites, and that the references triggered angry posts from followers, Heiss and Matthes surveyed followers and non-followers. Data indicated that respondents' anti-immigrant beliefs *drove* them to follow RWP sites that promoted anti-immigrant policies and provided a “community” of followers sharing such beliefs, resulting in a reciprocal “nativist spiral”: citizens with strong anti-immigrant attitudes exposed themselves to RWP content that reinforced their anti-immigrant attitudes.

Harlow and Kilgo (2021) showed how social media disrupt the “dominant media narrative” and “the hierarchy of social struggle” (Kilgo & Harlow, 2019) by amplifying protest coverage *differently* than mainstream news media do. Harlow and Kilgo sampled protest-related stories from national, metropolitan, and local newspapers, and collected Facebook “media engagement data” from its Application Programming Interface (API). Facebook users forwarded and redistributed news selectively, amplifying particular narratives and acting as “powerful gate-watchers” to mediate the news organization's role in legitimizing or delegitimizing movements (p. 680).

4 Introduction

Although these nine studies differ in purpose, focus, and techniques, they reflect a range of applications of *quantitative content analysis*—a research method defined briefly as *the systematic assignment of communication content to categories according to rules specified in a coding protocol, and the analysis of relationships involving those categories using statistical methods*.

Usually, content analysis involves drawing representative samples of content, training human coders to use a protocol to apply category rules to measure or reflect differences in content, and measuring the reliability (agreement or stability over time) of coders in applying the rules. Resulting data are usually analyzed to describe patterns or characteristics, or to identify relationships among the content qualities examined. If the categories and rules are sound and reliably applied, the chances are that the study results will be valid (e.g., that the observed patterns are meaningful). Though most of these procedures are well established, contemporary scholars are exploring new uses of computers to complement human coding and deal with large amounts of text, as discussed below.

This skeletal definition deliberately lacks mention of the researcher's specific goal (e.g., to test hypotheses about rap song lyrics), specification of types of communication to be examined (e.g., corporate websites, Instagram "selfies," or protest news), content qualities explored (e.g., presence of a reporter's opinion, a reference to mental health, or an anti-immigrant term), or types of inference that will be drawn from the data (e.g., that social media influence dominant narratives or that reporter and candidate agendas match). Such specification of terms is essential to an effective study design.

Moreover, the definition does not prescribe types of data analysis researchers might pursue. Some analyses are *univariate* (Vogt, 2005, p. 333), focusing on the distribution of cases on a single variable (e.g., the age distribution of models as child, teen, young-adult, adult, and elderly in television advertising), without relating it to any others. Other analyses are *bivariate* (p. 28), relating two variables (e.g., age distribution of models—the *dependent variable*—and type of programming airing the ads—the *independent variable*—to determine if elderly models, for example, are more often used in news than sports programming). *Multivariate* analyses (p. 201) involve relationships among three or more variables, often designating two or more independent variables (e.g., age distribution of models by programming type *and* by time of day—morning, afternoon, early evening, prime time). In later chapters, we emphasize that the complexity of data analysis is limited only by the research objective, how variables are measured (Chapter 5), and how the units of analysis are drawn (i.e., probability sampling; Chapters 6 and 9).

However, before a more comprehensive definition of this versatile method is developed in Chapter 2, we offer an overview of content analysis in mass communication research and examples of its use in other fields and disciplines.

Communication Research

Whereas some scholars approach communication messages from perspectives associated with the humanities (e.g., as literature or art), others employ a social science approach based in empirical observation and measurement. Typically, the latter approach means researchers identify questions or problems (e.g., derived from the scholarly literature or professional practices), identify concepts or factors that “in theory” may be involved, and propose possible explanations or relationships among concepts. Implausible explanations are discarded, and viable ones tested empirically, with theoretical concepts now measured in concrete observable terms.

If members of an ethnic minority, for example, believe that they are underrepresented in news content (in terms of their census numbers), a researcher may propose that it is because minorities are underrepresented among occupational groups that serve more often as news sources. This proposition, suggesting different concepts to be “operationalized” into measurement procedures, can be tested empirically. Similarly, if researchers want to address how social media fostered concerted activity during the 2020 Black Lives Matter protests, operational procedures can be developed and used to collect data on social media content, which might be compared with the protest activities of individuals who maintain the social media accounts.

Put another way, explanations for problems or questions for such researchers are sought and derived through direct and objective observation and measurement rather than through one’s reasoning, intuition, faith, ideology, or conviction. In short, these researchers employ what is traditionally referred to as the “scientific method.” The centuries-old distinction between *idealism* (i.e., the mind and its ideas are “the ultimate source and criteria of knowledge”) and *empiricism* (i.e., observation and experimentation yield knowledge) continues to hold the attention of those interested in epistemology or the study of knowledge (Vogt, 2005, pp. 105–106, 149). Content analysis assumes an empirical approach—a point made more emphatically in later chapters.

Content Analysis and Mass Communication Effects Research

Scholarly or scientific study of mass communication is fairly new, with roots in early 20th-century work by political scientists concerned with the effects of propaganda and other persuasive messages (McLeod, Kosicki, & McLeod, 2009). In addition to communication scholars, researchers from disciplines such as sociology, economics, and psychology have focused on communication processes and effects, contributing their own theoretical perspectives and research methods.

Powerful Effects?

One particularly durable communication research perspective reflects a behavioral science orientation that grew from early 20th-century theories that animal and human behaviors could be seen as stimulus–response complexes. Some communication researchers have viewed communication messages and their assumed effects from this same perspective.

Researchers interested in these effects typically adopted experimentation for testing hypotheses. Participants were assigned to different groups; some were exposed to a stimulus within a treatment (a message), whereas others were not (the control participants). Under tightly controlled conditions, subsequent differences in what was measured (e.g., attitudes about an issue or behavioral intention) could be attributed to exposure/non-exposure differences.

Meanwhile, for most of the first half of the 20th century, there was a widespread assumption—among scientists and the public—that stimuli such as mass persuasive messages could elicit powerful responses, even outside the experimental laboratory. Why?

Propaganda, as seen during the World Wars, was new and frightening (Lasswell, 1927; Shils & Janowitz, 1948). A 10-volume summary of 13 Payne Fund Studies conducted from 1929 to 1932 suggested movies' power to affect children's attitudes, emotions, moral standards, and perceptions of daily conduct (Lowery & DeFleur, 1995, p. 51). Anecdotal evidence of the impact of Communist or Nazi oratory in Europe, or the radio demagoguery of Father Charles E. Coughlin in America (Stegner, 1949), heightened concern over mass messages and collective behavior. Media were able to leapfrog official national boundaries and were believed capable of undermining national goals (Altschull, 1995). Broadcast media demonstrated a capacity for captivating, mesmerizing, and holding people in rapt attention, and inciting collective panic (Cantril, Gaudet, & Hertzog, 1940). With the rise of commercial advertising and public relations agencies, persuasive campaigns used messages crafted to make people do what a communicator wanted (McLeod et al., 2009).

These assumptions about powerful effects contributed to early models or theories that used metaphors such as *hypodermic needle* or *bullet*. In the language of the latter, all one had to do was shoot a persuasive message (a bullet) at the helpless and homogeneous mass audience, and the desired effects would occur. Experimental studies of messages and their effects were interpreted as supporting these assumptions.

Of course, the assumption that audience members were uniformly helpless and passive was a major one. Methodologists warned that the artificially controlled and contrived conditions in laboratory settings meant experimental attitude-change findings lacked real-world generalizability (Hovland, 1959). Others suggested that scientists' emphasis on how best to do things *to* the

audience was inappropriate; Bauer (1964, p. 322) questioned the “moral asymmetry” of such a view of the public.

Nonetheless, content analysis was useful within the powerful effects perspective because of the implicit causal role for communication content. It was important to study because it was believed to have an effect (Krippendorff, 2004a). Scholars scrutinized content for variables that could affect people. One might catalogue appeals used in propaganda, another might describe status differences among sources in persuasive messages, and a third might analyze whether anti-social behavior was sanctioned or ignored in television programs.

Limited Effects?

However, assumptions that powerful effects were direct and uniform were eventually challenged as simplistic (Severin & Tankard, 2000). Experimental findings had, in some cases, suggested that messages might change subjects’ knowledge but not the targeted attitudes or behaviors. Researchers conducting public opinion surveys brought field observations that ran counter to cause–effect relations found in laboratory settings.

Examination of how people are exposed to messages in the real world and mixed results on the effectiveness of persuasive message “bullets” suggested that a limited effects perspective might be worth exploring (Chaffee & Hochheimer, 1985). Non-laboratory audiences had only an *opportunity* for exposure to particular content; they were not forced to attend like experimental participants. Under “natural” conditions, audiences (who, surveys showed, were not uniformly helpless or passive, nor, for that matter, very uniform in general) used media and messages for their own individual purposes, chose what parts of messages—if any—to attend or ignore, and rejected much that was inconsistent with their attitudes, beliefs, and values (Lazarsfeld, Berelson, & Gaudet, 1944). A decision to accept, adopt, or learn a message was a function of existing psychological and social characteristics and not necessarily mere exposure to the manipulated, artificial credibility of a source trying to persuade. Social affiliations such as family and community involvement were important predictors of people’s attitudes and behaviors, and networks of personal influence were key in their decisions (Carey, 1996).

Contingency Effects?

Research during the second half of the 20th century thus suggested that the effects—powerful or limited—of mass media are contingent on a variety of factors and conditions. This contingency effects approach allowed theorists to reconcile conflicting conclusions of the powerful and limited effects approaches.

Rather than being the result of any single cause (e.g., the message), communication effects reflected a variety of contingent conditions (e.g., whether the message is attended to alone or as part of a group and what motivates one to attend). Put another way, the effect of a particular message may be *moderated* (Vogt, 2005, pp. 103, 195), or modified by attributes of individuals (e.g., horror movies are more frightening to young children; age is a moderator because it interacts with exposure). Exposure may be *mediated* through an intervening variable (p. 190): for example, exposure to footage of police using force during an arrest may have a different effect on a viewer whose father works in law enforcement than on a viewer whose father does not. The father's occupation intervenes and mediates the effect.

However, despite increasing interest in contingent conditions, in what people do with media messages, and in how—or if—they learn from them, content analysis remained an important means of categorizing content. Messages were now analyzed in terms of differences in psychological or social gratifications consumers might seek (e.g., escape from boredom, being “connected,” or having something to talk about), cognitive images they develop (e.g., views of gender roles or of the acceptability of antisocial acts), and what they deem important on the news media agenda (e.g., what issues in a political campaign were worth considering and what attributes of issues were critical). These studies of cognitive (not attitudinal) effects and people's social and psychological uses and gratifications of media and content reflected a view of the audience far different from the “morally asymmetrical” view criticized by Bauer (1964, p. 322). These triggered additional studies aimed at measuring content variables associated with those uses and effects.

For example, content analysts have categorized entertainment content to answer questions about how ethnic and gender stereotypes are learned (Mastro, 2009; Smith & Granados, 2009). They have looked at content ranging from daytime soap operas to reality programs because of guiding assumptions about psychological and social gratifications people achieve by viewing those shows (Rubin, 2009). They have examined victim gender in “slasher” horror movies because of concern that such violence is desensitizing (Sapolsky, Molitor, & Luque, 2003; Sparks, Sparks, & Sparks, 2009). And content analysis has shown how different communicators “frame” the same events, because scholars argue that frames shape interpretations (Biswas et al., 2021; Reese, Gandy, & Grant, 2001). According to Tankard (2001, pp. 100–101), “A frame is a central organizing idea for news content that supplies a context and suggests what the issue is through the use of selection, emphasis, exclusion, and elaboration.”

Moreover, as public and personal social media platforms have enabled virtually anyone to communicate publicly to audiences of unknown size, content analysis remains key to the study of those communicators and content. These

messages are not the work of traditional professional communicators, but they reach large audiences nonetheless, and scholars ponder the motives and gratifications of those consuming them and effects on their knowledge and beliefs.

Content analysis remains important for researchers exploring how individual-level cognitive processes and effects relate to message characteristics (Shrum, 2009; Oliver & Krakowiak, 2009). For example, scholars have argued that important differences between one message's effects and another's may be due less to the communicator's or audience member's intent (e.g., to inform or be informed) than to different cognitive or other processes (e.g., transportation and enjoyment, entertainment, arousal, mood management, social isolation, and so on) triggered by content features or structure (Bryant, Roskos-Ewoldsen, & Cantor, 2003; Green, Brock, & Kaufman, 2004; Oliver & Krakowiak, 2009; Vorderer & Hartmann, 2009).

These additional layers of complexity, and compelling questions of "what causes what" and whether relationships are unidirectional or even reciprocal (two variables mutually influence each other), point to the need—in *all* of social science, not just communication—for more sophisticated research designs that may incorporate multiple methods and data forms. Multi-method studies in communication might couple content analysis with surveys or experiments, while multi-form designs might use official transcripts as a baseline for comparison with mediated reports, or mainstream media coverage to compare with the responses and reinterpretations of online "gatewatchers" (e.g., Harlow & Kilgo, 2021), to name just a few examples.

Content Analysis and the Context of Production

Thus far, the discussion has implicitly viewed communication content as an antecedent condition, presenting possible consequences of exposure ranging from attitude change to the different gratifications people obtain from media or cognitive images they learn. However, content—whatever the medium that conveys it—is itself the consequence of a variety of other antecedent conditions or processes that may have led to or shaped its construction.

A news site's or aggregator's content, for example, is a consequence of the organization's selection from an array of possible stories, graphics, interactive features or affordances, and other content. That content may be a consequence of editors' application of what has traditionally been called "news judgment," based on numerous factors that visitors to the site need or want. The content is also shaped by other constraints, such as the kinds of motion graphics or interactivity available, how often material is updated, and so on. The content a researcher examines thus reflects all those antecedent choices, conditions, constraints, or processes (Stempel, 1985). In some instances, aggregator content is the consequence of an algorithm reflecting a user's prior choices.

Similarly, individual news stories are the consequence of influences including (but not limited to) a news organization's market (Lacy et al., 2010; Lacy, Watson, & Riffe, 2011; Lacy, 1987); resources available for staffing (Lacy et al., 2012; Fico & Drager, 2001); on-scene reporter judgments and interactions with purposive and non-purposive sources (Bennett, 1990; Duffy & Williams, 2011; Lawrence, 2010); and decisions about presentation style, structure, emphasis (as in the "framing" process described previously), and language, to name a few (Scheufele & Scheufele, 2010). Media sociologists no longer view news reporting as "mirroring" reality but speak instead of journalistic practices and decisions that constitute the manufacturing of news (Cohen & Young, 1981). News content is the product or consequence of those routines, practices, and values (Shoemaker & Reese, 1996; Reese, 2011), is constructed by news workers (Bantz, McCorkle, & Baade, 1997), and reflects both the professional culture of journalism and the larger society (Berkowitz, 2011).

Examples of "content as consequence" abound. Under the stress of natural disasters (e.g., tornadoes, hurricanes, or earthquakes), individual journalists produce messages in ways that differ from routine news work (Dill & Wu, 2009; Fontenot, Boyle, & Gallagher, 2009). Different ownership, management, operating, or competitive situations have consequences; news organizations in different competitive situations allocate content differently (Lacy et al., 2012; Beam, 2003; Lacy, 1992). The presence of women in top editorial positions has consequences for how reporters are assigned beats (Craft & Wanta, 2004) and the newsroom culture (Everbach, 2005), though evidence on the effects of female management on content is mixed (Beam & Di Cicco, 2010; Everbach, 2005) or perhaps issue-dependent (Correa & Harp, 2011). Predictably, some international coverage in US news media is a consequence of having a US military presence overseas; absent a state of war, "foreign news" is relatively rare (Allen & Hamilton, 2010). Facing censorship in authoritarian countries, correspondents gather and report news in ways that enable them to get stories out despite official threats, sanctions, or barriers (Riffe, 1984, 1991; Riffe, Kim, & Sobel, 2018). Symbols that show up in media messages at particular points in time (e.g., allusions to nationalism or solidarity during a war) are consequences of the dominant culture and ideology (Shoemaker & Reese, 1996); images, ideas, or themes reflect important antecedent cultural values.

"Content as consequence" is applicable to non-news communication, too. Recall Ki and Hon (2006), whose examination of Fortune 500 companies' websites allowed them to critique those companies' communication strategies—strategies that were antecedent to the site content.

Scholars often speak of such evidence as unobtrusive or non-reactive. That is, researchers can examine content after the fact of its production and draw inferences about the conditions of its production without making the communicators self-conscious or reactive to being observed while producing it (Weber, 1990).

Letters, diaries, bills of sale, or archived newspapers, tweets, or blog posts—to name a few—can be examined and conclusions drawn about what was happening at the time of their production, or what the producer *wanted* to have known about their production.

Indeed, with platforms and applications like Instagram, Twitter, Tumblr, and Facebook, content is a consequence of “construction” by individual users, for many reasons—to comment on events, signal support for a movement, share one’s experiences, or project an image that may or may not be accurate, to name a few. Original posts on such sites and comments are constructed content. Even the act of liking, linking, or forwarding material on these sites is an act of construction and communication.

The “Centrality” of Content

So, communication content may be viewed as end product, the assumed consequence of antecedent individual, organizational, social, and other contexts. The validity of that assumption depends on how closely the content evidence can be linked empirically (through observation) or theoretically to that context. As noted earlier, communication content also merits systematic examination because of its assumed role as cause or antecedent of a variety of individual processes, effects, or uses people make of it.

Figure 1.1 is a content-centered model illustrating why content analysis can be integral to theory-building about both communication effects and processes. The centrality remains regardless of the importance (for theory-building) of myriad non-content variables, such as individual human psychological or social factors and the larger social, cultural, historical, political, or economic context of communication.

However, if the model graphically illustrates the centrality of content, it does not accurately reflect the design of many mass communication studies. As Shoemaker and Reese (1990, p. 649) observed, most content analyses are not linked “in any systematic way to either the forces that created the content or to its effects.” As a result, Shoemaker and Reese (1996, p. 258) warned mass communication theory development could remain “stuck on a plateau” until that integration occurs. A 1996 study (Riffe & Freitag, 1997) of 25 years of content analyses published in *Journalism & Mass Communication Quarterly* revealed that 72% of the 486 studies lacked a theoretical framework linking content to either its antecedents or its consequences. Trumbo (2004, p. 426) placed the percentage at 73% in his analysis of *Journalism & Mass Communication Quarterly* content studies during the 1990 to 2000 period. Not surprisingly, only 46% of the cases examined by Riffe and Freitag involved formal research questions or hypotheses about testable relations among variables—testing that is essential to theory-building.

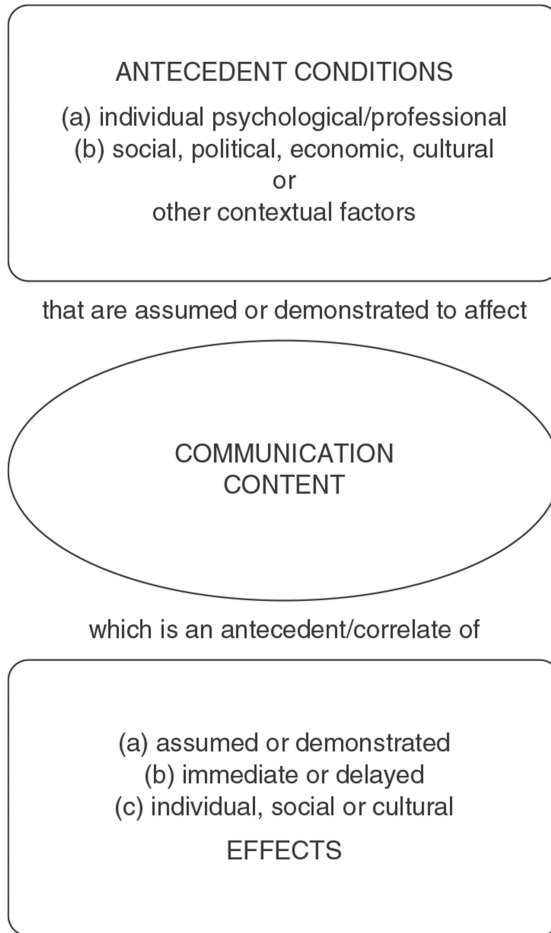


Figure 1.1 Centrality model of communication content

Still, research in this field is dynamic, although the scientific goal of prediction, explanation, and control (Reynolds, 1971) of media phenomena may still be decades away. However, quantitative content analysis of media content is key to such a goal. Since initial publication of this book in 1998, hundreds of content analysis-related studies have been published in *Journalism & Mass Communication Quarterly* and other refereed journals, such as the *Journal of Broadcasting & Electronic Media* and *Mass Communication and Society*, using the kind of quantitative content analysis examined in this book. According to Wimmer and Dominick (2011, p. 156), about a third of all articles published in those three journals in 2007 and 2008 employed quantitative content analysis, a higher

proportion than the 25% that Riffe and Freitag (1997) reported for 25 years of *Journalism & Mass Communication Quarterly*. Of the 2,534 articles Lovejoy, Watson, Lacy, and Riffe (2014) studied from *Journalism & Mass Communication Quarterly*, *Journal of Communication*, and *Communication Monographs* between 1985 and 2010, 23% involved content analysis.

Consistent with the emphasis on the “centrality” of content in understanding communication processes and effects, many studies place content analysis research into the context of framing, agenda-setting, cultivation, and various persuasion theories. Research on content antecedents is still largely atheoretical, though, with some studies using the Shoemaker and Reese (1996) hierarchy of influences approach to order, interpret, and interrelate influences on content. Theories addressing effects and antecedents of social media content will be difficult to synthesize, as such content ranges from comments on public events to Instagram selfies and school shooter manifestos.

Description as a Goal

Of course, not all research has theory-building as a goal. Simple descriptive studies of content have value. A Southern daily newspaper publisher, stung by criticism that coverage of the African American community was excessively negative, commissioned one of the authors to examine that coverage. The publisher needed an accurate description of his paper’s coverage to respond to the criticism and, perhaps, change the coverage.

Thus, some descriptive content analyses may be “reality checks” whereby representations or portrayals of groups, phenomena, traits, or characteristics are assessed against a “standard” (Wimmer & Dominick, 2011, pp. 158–159). Such comparisons to normative data can, in some instances, index media distortion (Mastro & Greenberg, 2000; Smith & Granados, 2009). For example, more than 30 years ago, a study of characters in television advertising during Saturday morning children’s programming reported a female and ethnic presence far smaller than those groups’ presence in the population, according to census data (Riffe, Goldson, Saxton, & Yu, 1989).

Historically, when new content and delivery forms evolve, they lend themselves to such descriptive “real-world” comparisons. Early video games, for example, were criticized because of assumptions about imitative aggression or learning of gender roles among users—a research focus previously applied to content ranging from comic books to movies, television, and popular music. Martins, Williams, Harrison, and Ratan (2008) analyzed 150 top-selling video games, measuring physical dimensions of animated characters and converting the dimensions to real-human “equivalencies.” Animated female characters were far more slender than their real-world counterparts—a pattern consistent with the thinness ideal cultivated by many media.

Or consider the study by Law and Labre (2002) analyzing male body images in magazines. Although the research used a longitudinal (1967–1997) design, it was essentially a descriptive study of how male body shapes became increasingly lean and muscular in visual representations. Law and Labre suggested that males' exposure to idealized mediated body images may parallel the experience women face.

Content on some social media applications and websites cannot always be assessed against a normative “gold standard” such as the census; instead, the content unabashedly reflects the authors' *own* world views or beliefs. Recall, for example, Heiss and Matthes's (2020) documentation of anti-immigrant and elite sentiments expressed on right-wing populist websites. Indeed, studies of sites that disseminate misinformation or blatantly false conspiracies require assessment only against what extant evidence and consensus confirm.

Finally, descriptive content analyses sometimes serve as a first phase in programs of research. Research on anonymous news sources is illustrative. Reporters sometimes hide a source's identity (e.g., “a senior official, speaking on condition of anonymity, said . . .”), despite complaints about the source's lack of public accountability (Duffy & Williams, 2011; Sobel & Riffe, 2016). Initially, Culbertson (1975, 1978) analyzed representative content to describe message variables associated with unnamed sources. Based on those results, Culbertson and Somerick (1976, 1977) conducted an experiment (participants received simulated news stories either with or without anonymous sources) to test the effects of unnamed sources on believability.

More recently, a program of research used experiments to test the effects of media framing of government policies on audience members, usually fashioning (manipulated) experimental treatment frames from examples found in analysis of media content (e.g., de Vreese, 2004, p. 39; de Vreese, 2010; de Vreese & Boomgaarden, 2006).

Research Applications: Making the Connection

As many of the examples presented above have shown, content analysis is often an end in itself—a method to answer research questions about content. However, some of the examples featured designs that brought together *multiple* forms or sources of content: Harlow and Kilgo (2021) studied news coverage and Facebook forwards; Lee and Riffe (2017) used press releases, corporate rating reports, and news coverage; and Bastien (2018) used newspaper coverage and debate transcripts. Other examples illustrate the method's use in conjunction with other methods: for example, Heiss and Matthes's (2020) analysis of right-wing populist Facebook posts and subsequent survey of followers and non-followers. In fact, numerous studies have involved multiple methods or data forms.

Scheufele, Haas, and Brosius (2011) explored the “mirror or mold” role of stock price and trading coverage on subsequent market activity. Data on four