Smith, Ruth Ann, and David S. Lux (1993), "Historical Method in Consumer Research: Developing Causal Explanations of Change." *Journal of Consumer Research* 19 (March), 595–610.

Song, X. Michael, and Mark E. Parry (1997a), "The Determinants of Japanese New Product Successes." *Journal of Marketing Research* 34 (February), 64–76.

______, and ______ (1997b), "A Cross-National Comparative Study of New Product Development Processes: Japan and the United States." *Journal of Marketing* 61 (April), 1–18.

Sorescu, Alina B., Rajesh K. Chandy, and Jaideep C. Prabhu (2003), "Sources and Financial Consequences of Radical Innovation: Insights from Pharmaceuticals." *Journal of Marketing* 67 (October), 82–102.

Szymanski, David M., Sundar G. Bharadwaj, and P. Rajan Varadarajan (1993), "Standardization versus Adaptation of International Marketing Strategy: An Empirical Investigation." *Journal of Marketing* 57 (October), 1–17.

Tellis, Gerard J., Rajesh K. Chandy, and David S. Ackerman (1999), "In Search of Diversity: The Record of Major Marketing Journals." *Journal of Marketing Research* 36 (February), 120–31.

______, and Peter N. Golder (1996), "First to Market, First to Fail? Real Causes of Enduring Market Leadership." *Sloan Management Review* 37 (Winter), 65–75.

______, and ______ (2002), *Will and Vision: How Latecomers Grow to Dominate Markets*. New York, N.Y.: McGraw-Hill.

______, Stefan Stremersch, and Eden Yin (2003), "The International Takeoff of New Products: The Role of Economics, Culture and Country Innovativeness." *Marketing Science* 22 (2), 188–208.

Utterback, James M. (1997), *Mastering the Dynamics of Innovation*. Boston, Mass.: Harvard Business School Press.

Webster, Frederick E., Jr. (1992), "The Changing Role of Marketing in the Corporation." *Journal of Marketing* 56 (October), 1–17.

Weiss, Allen M., and Erin Anderson (1992), "Converting from Independent to Employee Salesforces: The Role of Perceived Switching Costs." *Journal of Marketing Research* 29 (February), 101–15.

______, and Jan B. Heide (1993), "The Nature of Organizational Search in High Technology Markets." *Journal of Marketing Research* 30 (May), 220–33.

Wind, Jerry, and Vijay Mahajan (1997), "Issue and Opportunities in New Product Develop-ment: An Introduction to the Special Issue." *Journal of Marketing Research* 34 (February), 1–12.

Winer, Russell S. (1999), "Experimentation in the 21st Century: The Importance of External Validity." *Journal of the Academy of Marketing Science* 27 (3), 349–58.

9

Text Analysis as a Tool for Assessing Marketing Strategy Performance

José Antonio Rosa, Jelena Spanjol, and Joseph F. Porac

Introduction

Text analysis, also known as text mining, can be a powerful tool for assessing marketing strategy performance, particularly in the information-intensive 21st century. In contrast to data mining, which starts with small units of information (e.g., transaction information, customer characteristics) and aggregates them to generate a more complete picture, text analysis begins with complex self-contained documents and extracts bits of strategically important information by applying linguistic processing and classification techniques (Sullivan 2001). Academicians and practitioners use text analysis to distill important concepts from those extracted insights and to track underlying trends in market thinking and behavior from the overwhelming volume of text data that information technologies make available. For example, to gain a better understanding of competitors, customers, and the ambient technological and regulatory environment, companies such as Microsoft spend millions to maintain document warehouses and submit their contents to text analysis techniques. More recently, world-class market research firms such as IMS Health Inc. and NOP World have added text-analytic services (American Marketing Association 2003) to help their clients gain a finer appreciation of their customers. Researchers in marketing, strategic management, and information systems are using text analysis to study and better understand social and psychological phenomena important to marketing, such as how technologies are adopted across markets, how consumers identify with brands, and the competitive structure of markets. In every instance, text analysis involves assembling documents from multiple sources, classifying and marking them for systematic access and analysis, and clustering documents and analysis outputs according to topics that are compatible with the research being conducted.

Despite recent developments, we believe that the potential of text analysis as a tool for assessing marketing strategy performance has not been fully realized for two reasons. First, many researchers believe that text analysis is too expensive or difficult for them to use and hence do not implement it. Although text analysis at the levels described above is almost exclusively practiced by companies that can afford applications such as Oracle interMedia Text, LexiQuest Mine, and the IBM Intelligent Miner for Text, hardware and software advances are making much of that functionality increasingly accessible at a fraction of the cost. Moreover, object-oriented programming languages and browser-based user interface designs make modern text analysis applications easy to use, and have reduced the time required to achieve operational proficiency. One objective of this chapter is to help dispel the myth that only Fortune 500 firms and well-heeled technology-savvy researchers at major universities can afford the tools required for effective text analysis.

The second reason that text analysis is underutilized is that even those currently conducting text analyses have only a limited vision of what can be gleaned from text sources. Currently, text analysis is used primarily to qualitatively supplement the more traditional quantitative measurement of important variables, be they consumer attitudes, employee behaviors, competitor tactics, or market trends. Most researchers overlook the stand-alone value of text analysis for clarifying and quantifying social forces that shape organizations, markets, and environments, but that are not otherwise discernible, and for identifying likely future shifts in consumer and competitor behaviors. Sensemaking—deriving rationalizations for beliefs and behaviors—is a powerful organizing force in business and society at large (Weick 1995), but it is seldom if ever captured by demographic trends, advanced orders, scanner data, or even by the attitudes and beliefs of consumers and producers as measured by standard survey research. It *is* captured, however, by the narratives of those involved, and narratives are a type of data highly accessible through text analysis.

Throughout this chapter we focus our attention on one particular type of narrative: market stories. We explain the importance of market stories, describe

text analysis tools and techniques that can be applied to studying market stories, and provide some examples of how text analysis of market stories reveals a market's underlying trends and relationships and can therefore be used to assess the effectiveness of existing marketing strategies. We conclude by discussing future directions for text analysis in marketing strategy research.

What Are Market Stories and Why Are They Important?

Market stories (also called narratives) are used by market actors such as consumers and producers to describe the storyteller, people, or events she or he has observed, and other aspects of the market environment in terms of attributes and cause-effect relationships. Consider as an example the following excerpt from a news publication:

> The West Towne Mall in Madison, Wis., is bustling with young shoppers on Saturday afternoon, but at one retailer there's plenty of elbow room: Abercrombie & Fitch Co. As 1 p.m. approaches, only four customers are in the store, and two of them are heading out the door empty-handed. "You can buy two pair of jeans elsewhere for the $60 you pay here for one," gripes Chris Budzin, 20, as he makes his exit with girlfriend Jayne Vosen, 18. (Berner 2003, p. 90)

In this example we have a story within a story. First we have the news article, in which the journalist describes the situation within an Abercrombie & Fitch store and quotes a complaining customer as part of his own explanation for what is wrong with the company's current strategy—an explanation that the journalist develops in the rest of the article. The journalist, in other words, is telling a market story about Abercrombie & Fitch that explains the decline that he and possibly others have noticed.

In addition, we have the customer (Chris) telling his own market story as a way of making sense of why he is leaving the store. In truth, we don't really know if Chris is leaving because he cannot afford the clothes at Abercrombie & Fitch (he does not have $60), because he has recently become less willing to spend $60 on jeans than he used to be (even though he has $60, spending them on a pair of jeans from Abercrombie & Fitch is no longer a good idea); or because his girlfriend mentioned that she had recently purchased jeans elsewhere for $30 and he wants to come across as an equally wise shopper. And it is quite possible that Chris is not sure of his true motivations either. But to Abercrombie & Fitch marketing managers, and to researchers studying retailer strategies, what actually motivated Chris's exit should not matter much. What matters most is that he made a public sensemaking statement about Abercrombie & Fitch's prices being too high to someone else, that the statement found its way into the pages of a nationally distributed business publication, and that there now exist thousands or possibly millions of people who may come to believe that Abercrombie & Fitch products are over-priced, and who may then behave (e.g., shop, invest, recommend the stock to others) in ways compatible with that belief.

Research shows that market actors tell stories to make sense of their own behaviors and preexisting beliefs, to check if their thinking resonates with that of other people or companies in the same market arena, and possibly to influence the thinking of others. A good story holds disparate elements together long enough to energize and guide action, plausibly enough to allow people to make retrospective sense of whatever happens, and engagingly enough that others will contribute their own inputs in the interest of sensemaking (Weick 1995).

Market stories, in other words, stem from our need to understand why we behaved as we did, or why whatever just happened actually happened. Once the story is out, however, it invites others into the process of behavior and sensemaking, and in effect empowers the story to shape future actions. In the earlier example, for instance, Chris (the customer) may have initially had muddled reasons for leaving Abercrombie & Fitch. Once Chris shared his story with the journalist, however, high prices became more firmly entrenched in his own mind as the reason for his leaving, and overpriced goods became more strongly associated with Abercrombie & Fitch's image, both in his own mind and that of his audience. From a cognitivist perspective, market stories generate the knowledge structures that help market actors reconcile current behaviors with preexisting beliefs, and by doing so they also inform and coordinate future behaviors. That being true, it is not surprising that market stories can be used as windows into the future of product markets. In order to peer through these windows, however, stories need to be captured and analyzed, tasks greatly facilitated by available and emerging text analysis techniques and tools.

The role of market stories in the shaping of product markets has been firmly established by research in the sociology of markets and technologies (e.g., White 1981; Pinch and Bijker 1987; Garud and Rappa 1994; Bijker 1995) and by strategy and marketing researchers (e.g., Garud and Rappa 1994; Rosa et al. 1999). Some have gone so far as to argue that markets *are* conversations (Levine 2000). Even if markets are more than conversations, stories are essential to how we navigate and make sense of our world (e.g., Gabriel 1998; Weick 1995). Given their multiple functions in shaping our social realities, stories are critically important in developing and understanding marketing strategy.

How Do Stories Impact Product Markets?

Market actors, whether consumers, producers, retailers, or others, need stories to understand new and existing products; to know how to sell, use, or compete with such products; to assess whether or not products provide value; and to identify where to position new products relative to the many other artifacts that shape their lives. Automotive consumers, for example, use their own stories and those of others to make sense of new products relative to preexisting and evolving products. Consumers also rely on market stories to maintain in their minds how old and new products relate to each other; the stories provide a framework on which they rely every time they enter the marketplace. Likewise, competing producers use market stories to interpret and respond to new market entries and to make sense of new product-usage combinations (e.g., crossover vehicles as traveling offices) that consumers often enact in response to new product entries. Examples of how market stories have shaped product markets can be found in the histories of products as diverse as paper clips (Petroski 1993), bicycles (Bijker 1995), local-area network (LAN) technology (Theoharakis and Wong 2002), and cochlear implants (Garud and Rappa 1994), suggesting that market stories can influence a wide variety of product markets.

Who Tells Market Stories?

Some product market stories begin when actors within consumer or producer networks (e.g., dealers, journalists, opinion leaders) experience new products and summarize their experiences through dialogue with others. In our earlier example, the Abercrombie & Fitch customer told a story that explained why he was leaving the store empty-handed. Stories can also begin with journalists using market data and personal observations to explain industry shifts, with company representatives positioning their products and justifying their actions, and even with government employees explaining policy changes. Regardless of their origins, market stories are broadcast to other market actors as part of the sensemaking process, and stories that resonate with large audiences often find their way into published sources and are widely disseminated.

Research suggests that whereas all market actors need stories to navigate products markets, not all market actors are storytellers (e.g., Gladwell 2000). A case in point is described by Sirsi, Ward, and Reingen (1996), who find that not all market actors tell coherent stories about health foods. A minority of consumers hold sufficiently rich arrays of concepts and assumed causal relations (e.g., potatoes are nightshade foods, nightshade foods cause arthritis) to assemble them into compelling stories: "Potatoes should be avoided because they are nightshade foods, and nightshade foods cause arthritis." Others, who are less adept at storytelling, can articulate only portions of the concepts and relation arrays held by storytellers. When it comes to consumption behavior, however, storytellers and nonstorytellers behave in almost identical ways, suggesting that nonstorytellers use the narratives shared by storytellers to inform their own behaviors. Likewise among producers, Porac et al. (1995) found differences in the abilities of Scottish knitwear producers to articulate stories about their strategies. Some told cogent stories about market leaders and followers in different market segments and why the prevailing industry structure made sense, while others borrowed from those stories to shape their own strategies and experiences. In general, social actors need stories of identity—narratives that help them think about and feel who they are, where they come from, where they are headed, and why they do what they do (Gardner 1995; Gabriel 1998). Not all of them, however, need to formulate their own stories, and not all stories are equally good at generating and sustaining meaning.

How Are Stories Shared?

While the majority of market actors do not generate stories, they need access to the stories of others. Depending on the domain, however, market story sources can vary dramatically. In the case of consumer products, for example, stories may first start circulating at club and professional meetings, soccer games, and more recently, in Internet-based bulletin boards and chat rooms. Other stories start in, or eventually find their way into, industry journals, enthusiast publications, publications that claim to be objective protectors of consumer interests (e.g., *Consumer Reports*), and the topical columns (e.g., auto, travel, home decor, technology) in daily and weekly newspapers and news magazines. Particularly in the absence of one-on-one contact with storytellers, market actors turn to public and published sources. Stories both reflect and inform market behaviors, and they often lead to widespread behavioral trends. It is this dual function that makes them valuable windows into the future. Table 1 presents a list of factors that may be relevant to marketing strategy research and possible sources for related market stories. The list is far from exhaustive, but it may nevertheless enrich the thinking of the reader.

While we have focused our discussion on market stories, it is important to note that narratives in general can carry great strategic significance across the spectrum of management functions and organizational types. Stories are vital to all

Table 1
Influences on Marketing Strategy and Potential Sources of Relevant Stories

Influences on Marketing Strategy	Possible Sources of Text
Consumer word-of-mouth reactions to new products and other marketing efforts (price, locations, advertising campaigns, etc.)	Product enthusiast chat rooms and bulletin boards Letters to enthusiast magazines and publications Dialogue at enthusiast and user club meetings Customer correspondence Transcriptions of conversations on customer service hotlines
Consumer brand associations	Product enthusiast chat rooms and bulletin boards Customer correspondence Positioning statements used by grassroots enthusiast or detractor organizations for the focal brand and for competing brands
Competitor strategies	Competitor press releases Interviews in industry publications with senior managers from competing firms Competitor company forecasts and presentations to the investor community, letters to shareholders, mission statements Industry analyst reports on companies and the industry at large
Emerging technologies that may influence current products or generate new ones	Articles and presentations at professional meetings (e.g., Society of Automotive Engineers, American Medical Association, etc.) Conference proceedings, often posted on industry and association websites Professional publications in basic and applied sciences, often available online Patents records containing descriptions of emerging technologies
Government and social-policy initiatives that may influence current products or generate new ones	Government agency reports Chat rooms and websites for special interest groups Articles and presentations at special-interest group meetings Conference proceedings, often posted on special-interest group and association websites Industry analyst reports on the industry at large LexisNexis™ congressional and government periodical database

organizations (Weick 1995), and particularly when an organization is involved in developing and implementing new initiatives. More specifically,

> Stories create, sustain, fashion and test meanings in and out of organizations. They are part of a sense-making process which can be researched *in situ* . . . the truth of a story does not lie in *the facts*, but in the meaning. (Gabriel 1998, p. 85)

Within firms, for example, stories are used to initiate and enable organizational change by enhancing shared understanding and by facilitating the creation of compelling strategic plans (Denning 2001). They can also be used to generate shared knowledge in professional groups (e.g., physicians, attorneys, CPAs), in which uniformity of thought and behavior are important for legitimizing exclusivity and limiting membership, or to mobilize special-interest groups into collective action. Table 2 provides some examples of how text analysis has been used in marketing and other research areas. The list focuses only on published academic research, but companies and government agencies also use text analysis, as do consultants and market research firms, both for commercial and private research.

Methods, Tools, and Applications of Text Analysis

Text analysis can have a qualitative or quantitative orientation, and both approaches can be combined in single studies. Qualitatively oriented text analysis typically consists of free-form, manual coding of texts (with or without facilitating software) from which recurring themes are extracted and summarized. Quantitatively oriented analysis also relies on the coding of text, but is more focused on automated coding and concept occurrence counts and utilizes dictionaries (user-defined or existing) and co-occurrence or semantic network approaches (Roberts 2000; Züll and Landmann 2002). Depending on the research objectives, approaches can be combined to answer specific questions from the compiled text databank.[1] The motorcycle industry study described later in the chapter combined qualitative and quantitative approaches; it used automation for some of the text coding. The minivan market study (also described later) relied purely on manual coding, although co-occurrence counts were developed and used for quantitative analysis.

In general, researchers go through several steps when conducting text analysis. Figure 1 delineates the major stages that most researchers will encounter, using Rosa et al. (1999) as an illustrative example.

Table 2
Business Research Applications of Text Analysis

Author(s) and Study Year	Text Source Analyzed	Research Target
Ober et al. (1999)	Annual reports and other publicly distributed corporate documents	Identifying words that communicate the company's certainty about its future and the adequacy of its policies and strategies
Rosa et al. (1999)	Articles from consumer and industry publications pertaining to the minivan market	Tracking how consumers and producers contributed in defining the minivan concept in the early stages of the market
Rosa, Judson, and Porac (2005)	Articles from consumer publications in the motorcycle market	Ascertaining the mutual influence of industry category structure and product models on each other
Shaw, Brown, and Bromiley (1998)	Oral and written narratives comprising intra-organizational correspondence and documentation from meetings	Examining the dissemination of strategic objectives and initiatives throughout the organizations responsible for their attainment or implementation
Sydserff and Weetman (2002)	Statements by senior executives to investment analysts	Evaluating the relationship between the use of active or passive voice and the perceived effectiveness of the speaker as a communicator
Theoharakis and Wong (2002)	Abstracts from articles on local area network (LAN) technology from industry publications	Ascertaining the influence of public discourse on the success or failure of competing technologies
Venette, Sellnow, and Lang (2003)	Media reports and NHTSA committee hearings transcripts	Creating a metanarrative model of ways in which organizations can manage and influence public narratives surrounding organizational crises
Vaara (2002)	Transcribed interviews with managers, internal company documents, and news materials	Examining narratives of successful or unsuccessful organizational integration following mergers and acquisitions

As in any research project, regardless of the selected methodology, the first step involves defining the research question, with hypotheses when appropriate, and identifying a theoretically based rationale when the research is deductive. If the

Figure 1

The Process of Text Analysis Research: An Application to Rosa et al. (1999)

1. Theory, Research Questions, and Rationale
Questions: What are the research questions? What is the theoretical perspective taken? Are there hypotheses?
Rosa et al.: The research investigates how product markets function and evolve, whether market boundaries are distinct and stable or shifting and overlapping. A sociocognitive perspective is taken. Hypotheses include that there are differences in references to product category labels in market conversation over time and that membership of products in product categories depends on attributes and category stability.

↓

2. Variable Definitions and Operationalizations; Sampling of Texts
Questions: What are the variables of interest? How will they be measured? What is the unit of analysis? What are the text sources, and are texts sampled in their entirety or partially? What are the validity and reliability issues?
Rosa et al.: Variables of interest include frequency of product category references in texts (e.g., minivan, car, station wagon, van), and acceptability of different minivan models. Unit of analysis was one market story (i.e., published article). Articles from both producer and consumer publications were sampled in their entirety.

↓

3. Development of Coding Categories, Rules, and Process
Questions: Is a coding dictionary available, or does it need to be developed? What are the coding categories and rules? Is coding automated, done manually, or a combination of both? What software is most appropriate for the coding procedures?
Rosa et al.: Coding categories and rules were developed for the study. Coding was done manually on ASCII text by two independent coders with the ATLAS.ti software package.

↓

4. Training and Coding
Questions: Who will be coding the data? If there will be multiple coders, will they require training? Is initial reliability established? Does the coding procedure (categories, rules) need revising?
Rosa et al.: Two independent coders coded all the data. Training was completed before main coding began, and necessary revisions to the coding procedure were made after satisfactory inter-rater and coding-scheme reliability were established.

↓

5. Analysis and Reporting
Questions: Is final reliability for inter-rater coding and variables coded appropriately? Is multivariate analysis of coding results appropriate, and what are the results?
Rosa et al.: Inter-rater reliability was measured with Cohen's Kappa and found to be significant. Analysis of coded data included Poisson maximum likelihood regression and [illegible]

Adapted from: Neuendorf, Kimberly A. (2002), *The Content Analysis Guidebook*. Thousand Oaks, Calif.: Sage Publications, 50-1.

research is inductive and theory is expected to emerge from the data, research questions can be more loosely defined and hypotheses are not developed a priori. In such cases the data must be analyzed carefully and iteratively, and insights are often restructured or redefined repeatedly. Inductive text analysis often results in coding schemes having to be discarded or significantly altered throughout the process.

Regardless of whether the research is inductive or deductive, research questions (and possibly hypotheses) become more specific as the researcher enters the second stage in text analysis, which is marked by the development of a set of concrete variables of interest, including conceptual and operational definitions. In defining the variables, researchers must give particular attention to measurement and analysis feasibility, which can be tricky with text data because of the ambiguities inherent in language. In addition, the availability and accessibility of text data come into central focus in this stage. Researchers must identify adequate sources, ones that offer the market stories of interest and preferably in digital formats compatible with the software to be used (more on this point in the description of software tools below). For market and competitive studies, publicly available news sources are often sufficient. The Rosa et al. (1999) study of the minivan market, for example, used consumer and industry publications to capture the market stories of both consumers and producers, while Rosa, Judson, and Porac's study in the motorcycle industry (2005) used only consumer publications. Other studies have relied on industry and government publications and correspondence (e.g., Garud and Rappa 1994; Venette, Sellnow, and Lang 2003) and transcripts of interviews (e.g., Vaara 2002). When studies are focused on intrafirm strategic decision making and implementation, accessing stories from internal memoranda, e-mails, operational policies, and interviews with key players can be challenging, but is seldom a terminal constraint. Some factors to consider when selecting sources are market maturity (is the market new or established?), the level of industry and market consolidation (are there many or few producers or customers?), and the reputation of the different sources of market stories, in addition to the overall number of sources available.

Following the preparation of the text database, researchers need to develop a coding scheme and the coding rules that will guide the people or automated procedures actually analyzing the text. Regardless of the topic area, researchers must define coding categories that are relevant to the research question, clear and unambiguous, complete in capturing the relevant themes occurring in the text, and mutually exclusive (Züll and Mohler 2001). For inductive research completeness is achieved iteratively, and its sufficiency is determined late in the research process. For deductive studies, these requirements should ideally be fulfilled before coding commences. Alternatively, a subset of the sample may be used to test out the coding scheme and adjust it as necessary to achieve adequate completeness.

Several text analysis packages offer already developed dictionaries, with word lists for different variables embedded in the program (see the description of DICTION later in the chapter). If such a dictionary-based approach is used, each coding category is defined by a particular word or phrase list (or both), and the coding can be automated, tracking the occurrence of focal concepts by source and unit of analysis (e.g., sentence, paragraph, number of lines). In a co-occurrence coding approach, such as was used in the motorcycle study, the text corpus is searched for the co-occurrence of specific words or phrases within the same unit of analysis, counting the instances of co-occurrence and possibly measuring the distance between concepts in terms of number of words involved. Alternatively, in semantic network analysis, the data are coded for specific interactions between terms or phrases, resulting in node-relation-node networks.

Once the coding rules have been developed, the actual coding is conducted either manually by individual coders or automated within the software (see Table 3 for a comparison of the capabilities of existing coding software). Factors to consider in the choice of coding methods can include time and cost requirements (e.g., training for individual coders), inter-rater reliability and interpretation issues, appropriateness of the coding scheme for automation, availability and selection of appropriate software tools, and the feasibility of articulating coding rules and dictionaries clearly enough for programmers to do their job. Once the coding is complete, the output can be statistically analyzed.

Researchers interested in text analysis often ask how long it takes. There are several important factors that influence how much time a text analysis project will require. The complexity of the research question and the number of hypothe-ses are directly and positively related to time required to complete a project. Simple research questions that require text to be coded for only two or three variables will likely involve only one pass through the text data and relatively little time for the coding. If the average person reads for comprehension at 150 words per minute, and the typical page of single-spaced text contains 600 words, it takes on average four minutes to read a page of text and another one or two minutes for simple cod-ing. Under such conditions, coding the 116 articles used in the Rosa et al. (1999) study can be estimated to have taken approximately 11 hours for each coder. In fact, however, the Rosa et al. (1999) study coded for hundreds of distinct concepts, and at multiple levels of abstraction, sometimes demanding 30-40 minutes per page of text. In the end it took over 720 person-hours to code the text used by Rosa et al. (1999), clearly illustrating that coding complexity is a major contributor to

Table 3
A Summary of Text Analysis Software Packages and Coding Capabilities

Program	Mode of Coding	Dictionary	Quantitative	Qualitative
AQUAD	Manual	Hierarchical		✓
ATLAS.ti	Manual, automatic in combination with search			✓
CoAn	Automatic, interactive	Word based	✓	
Code-A-Text	Manual, automatic	Word lists, nominal or ordinal scales		✓
DICTION	Automatic	Hard coded, word based	✓	
DIMAP-MCCA	Automatic	Hard coded, word based	✓	
Ethnograph	Manual			✓
Hyper RESEARCH	Manual, restricted automatic			✓
KEDS	Automatic, interactive	Domain specific, word and rule based	✓	
Kwalitan	Manual, automatic in combination with search			✓
NUD*IST	Manual, automatic in combination with search			✓
QED	Manual, automatic in combination with search			✓
TATOE	Manual, automatic in combination with search			✓
TextGrab/TextQuest	Automatic, interactive	Hard coded, word based, also user defined	✓	
TEXTPACK	Automatic	Word based	✓	
TextSmart	Automatic, interactive	Automatically generated or user-defined categories, word based	✓	
VBPro	Automatic	User defined, word based	✓	

Program	Mode of Coding	Dictionary	Quantitative	Qualitative
WinMaxPro	Manual, automatic in combination with search			✓
WordStat	Automatic, interactive	Word based	✓	
TextAnalyst	Automatic	Automatically generated or user-defined categories	✓	

Expanded and adapted from: Alexa, Melina, and Cornelia Züll (2000), "Text Analysis Software: Commonalities, Differences, and Limitations: The Results of a Review." *Quality and Quantity* 34 (3), 299-321.

the time required for text analysis research. Also contributing to the time required for text analysis is the total amount of text being studied, its readiness to be imported into text analysis software, and the ease with which coding rules can be explained to coders or translated into data handling rules for automated coding. Acquiring and preparing text documents for use with text analysis software can be time consuming, particularly if it involves scanning and character recognition of hard-copy documents. Also time consuming is developing coding rules and training coders (humans or machines) in their proper use. There is no easy-to-use rule of thumb for estimating time requirements for text analysis research, other than to say that in the same way that ethnographic research is more time consuming than survey research because of the complexity of extracting nuanced meanings from narratives, text analysis is likely to be more time consuming than other research approaches because it relies on the same ambiguous substrate (language) for its insights. For many questions in the field of marketing strategy, however, the time spent on text analysis is a sound investment given the rich insights that researchers can glean from studying narratives, as illustrated in the next section.

Using Text Analysis to Assess Marketing Strategy Performance

We now turn our attention to two examples that show how text analysis can be used to assess marketing strategy performance. The first example deals with emerging markets, the second with mature markets. The descriptions are abridged because of space considerations, but published papers are available for both studies.

Emerging Markets

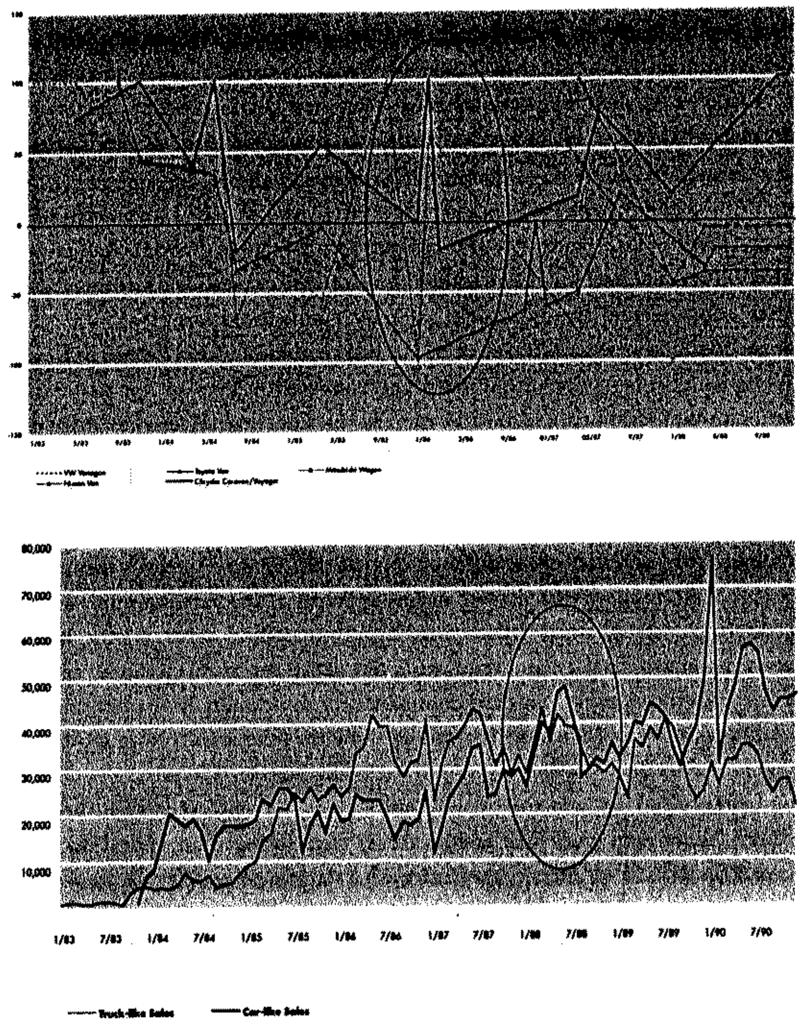
A longitudinal study of the minivan market (Rosa et al. 1999) found that during turbulent periods of market ferment, market stories can be a reliable source for

likely market trends. In addition, early signals communicating a producer's future strategy to competitors and consumers can influence market development. Market stories[2] representing consumer and producer views were compiled and coded for mentions of product attributes, such as front-wheel drive or seven-passenger seating, as well as for favorable and unfavorable mentions of different product models, the minivan category in general, and competing categories. The text analysis output was subsequently submitted to quantitative analysis using categorical time series regression to ascertain the relationship between mentions of product categories and models and variables of interest to marketing strategists, such as product model sales, market stability, and product model preferences among consumers.

As in most emerging markets, the early years of the U.S. minivan market were marked by little consensus over minivan architecture (e.g., front- or rear-wheel drive, passenger capacity, cargo configurations) and usage (e.g., commercial purposes, supplemental family vehicle, primary family transportation), and several domestic and import brands competed for market dominance. The research coded, among other things, the use of the term "minivan" relative to key architectural attributes that were being contested, such as the drive wheels, engine positioning, and carrying capacity. As part of the broader study, the coding of these data sought to identify the time frame within which market consensus emerged on what constituted a minivan (the product's common architecture), and the effects of that consensus on the popularity of the different models that claimed category membership.

Results show that market stories anticipated the car-like minivan architecture that persists today (characterized by front-wheel drive, engine in front, seven-passenger capacity, and room for a flat 4x8 sheet in cargo mode) two years before sales trends signaled a clear dominance for this design. In market stories circulating as early as 1986, concept counts show that the architecture most closely exemplified by the Chrysler minivan (the car-like Voyager/Caravan) was dominating market thinking, and that market entries not in compliance with that architecture (those with truck-like architectures) were losing their appeal, facts confirmed by Rosa et al.'s analysis of sales trends and new product announcements into the 1989 model year. Figure 2 illustrates how market stories were an effective window into future market preferences in the minivan market. Note that whereas the tone of market stories about the Chrysler minivan and other models suggested the Chrysler architecture was preferred around 1986 (the circled area in the upper diagram of Figure 2), the sales volumes of what came to be called car-like minivans did not come to dominate the industry until 1988 (the circled area in the lower diagram of Figure 2).

Figure 2
Minivan Model Acceptability Scores Drawn from Market Stories and Actual Minivan Sales Data (1983–1990)



Source: Rosa, José Antonio, Joseph F. Porac, Jelena Runser-Spanjol, and Michael S. Saxon (1999), "Sociocognitive Dynamics in a Product Market." *Journal of Marketing* 63 (Special Issue), 64-77.

Mature Markets

Text analysis can also be used to assess market strategy performance in mature markets, as illustrated by Rosa, Judson, and Porac (2005). This research focused on the relationship between product categories and product models in the motorcycle market. The authors proposed that producer and consumer behaviors and sense-making in the fast-moving motorcycle market give rise to a mutual dependency between product models and product categories. That is, producers and consumers understand product models in terms of the product categories with which the models are associated, and those categories, in turn, are defined by the models associated with them.

In effect, the Rosa, Judson, and Porac (2005) study found that the actions of market actors often have direct bearing on the category structures that shape markets, and that in positioning new and existing product models, producers can reinforce or undermine, sometimes unknowingly, those elements of the category structure that make profitable exchange possible. The study's results show that the number of product models attached to a particular category label influences its longevity, expansion, or decline as an effective positioning anchor for new models. Even more important to marketing strategy, the results also show that the longevity of specific models after their introduction is influenced both by how strongly they are conceptually bound to specific category labels and by the number of category labels to which the model is linked. Of particular importance was the finding that, at least in the motorcycle industry, new product models benefited from having strong associations with more than one product category, quite possibly because such associations gave the models access to a higher number of beliefs and attributions among motorcycle enthusiasts. Multiple category associations seem to provide motorcycle models with a more diversified conceptual anchoring and to protect them from category label drift, a common phenomenon in fast-moving markets. This is an important outcome when product development costs millions of dollars and requires more than one product-year's worth of sales to recover the investment.

As in Rosa et al. (1999), a library of relevant market stories was compiled, consisting of the complete text of all *Cycle World* product review and editorial articles between 1990 and 1996 in digital format.[3] The relationship between category membership and model longevity was tested by regressing model longevity against the number of categories with which the model was associated and the strength of association with the categories. Association strength was operationalized as the proportion of individual product model mentions in market stories that were linked to each specific category. Model longevity was operationalized as the total number of consecutive years that a model was mentioned in the text corpus. All coding of the *Cycle World* data was conducted with the content-analytic software package ATLAS.ti (Muhr 1997).

Tools for Text Analysis Research

ATLAS.ti is only one of the many text analysis tools available to researchers. In this section, we highlight and compare several highly differentiated tools for text analysis and suggest potential marketing strategy research uses for each one. We focus on four maximally different software packages in order to give readers a sense of the variety of approaches and tools available. The dimensions on which these software tools differ are coding scheme flexibility, level of coding automation, text acquisition and preparation capabilities, and qualitative and quantitative analysis integration and exploratory capabilities. Table 3 gives an overview of these and additional text-analytic software packages, but it is not an exhaustive list of available software. New products are introduced often, and researchers are advised to check some of the websites listed in the appendix for information on newly available software and upgrades.

ATLAS.ti

A Windows-based software package, ATLAS.ti (Muhr 1997) allows users to assemble and organize large databases of text, pictorial, and sound documents for the creation of coding schemes that cut across documents. The software also allows for extensive annotation of focal documents and codes in multiple formats, aggregation of codes and documents into hierarchical families, rudimentary automated searches, and the creation and updating of graphical representations of conceptual networks as they emerge from the data. ATLAS.ti users have considerable flexibility in devising a context-specific coding scheme, which makes the software package extremely pliable. All of the text analyses done in the minivan (Rosa et al. 1999) and motorcycle (Rosa, Judson, and Porac 2005) studies were conducted by individual coders working on personal computers using ATLAS.ti, and both were completed in a matter of weeks. Furthermore, both projects yielded in-depth and quantifiable representations of product-market phenomena that could have been easily maintained as more years of data were accumulated.[4]

One drawback of ATLAS.ti and similar software is the effort required to prepare documents for input and processing. As with similar packages, ATLAS.ti works best when ASCII text files are used. Although still the most common format for computer text files, ASCII has been replaced by Unicode in newer operating

systems (such as Windows 2000), and given the predominant use of word processing, HTML, or portable digital (PDF) formats in the archiving of available text, ATLAS.ti's preference for ASCII files often requires researchers to perform additional steps when assembling a document database. PDF files, for example, have to be filtered through character recognition software to create ASCII files, while word processing or HTML documents need to be reformatted as text files with carriage returns after each line. Extraneous punctuation marks and other text delimiters often have to be stripped from the text before analysis. More costly (or specialized) text analysis software, some of which we discuss below, can be more ecumenical in its acceptance of document formats and may be a better choice for researchers who have sufficient financial resources to purchase specialized software.

DICTION

In contrast to ATLAS.ti's coding flexibility, DICTION (Hart 2000) carries a hard-coded dictionary and offers entirely automated coding and processing capabilities. Also a Windows-based program, DICTION analyzes the tone of a verbal message by searching text units for 5 general features and 35 sub-features. The 5 general features sought are (Hart 2000):

- Certainty: Language indicating resoluteness, inflexibility, completeness, and a tendency to speak ex-cathedra
- Activity: Language featuring movement, change, the implementation of ideas, and the avoidance of inertia
- Optimism: Language endorsing some person, group, concept, or event or highlighting their positive entailments
- Commonality: Language highlighting the agreed-upon values of a group and rejecting idiosyncratic modes of engagement
- Realism: Language describing tangible, immediate, recognizable matters that affect people's everyday lives

Whereas context-derived coding can face significant problems in terms of validity and reliability (especially when multiple coders are involved), a hard-coded dictionary offers ease of use, objectivity, and reliability. Contributing to DICTION's reliability is the impressive size of its dictionary (10,000 words), its systemic approach to language study, and the fact that the dictionary was constructed from the analysis of more than 20,000 texts and includes standard scores on all variables in different contexts (such as politics, business, and so on). The verbal tone variables can be used as predictors by themselves or in conjunction with other coding schemes developed as custom dictionaries within DICTION or coding schemes in other software packages. All output from DICTION is assembled into numeric files that can be easily transported into statistical analysis packages such as SPSS and SAS.

DICTION has been utilized in the analysis of strategic accounting narratives (e.g., Ober et al. 1999; Sydserff and Weetman 2002), with findings suggesting that market stories and strategic narratives contain information above and beyond the surface-level information. DICTION analysis goes beyond the expressed informational content of text data to focus on hidden meanings embedded in the syntactic structure of the text. Sydserff and Weetman (2002), for example, examined chairmen's statements and managers' reports from investment trust companies and captured how "good" performers communicate differently from "poor" performers through these narratives. They found that poor performers communicate in a more passive tone. Sydserff and Weetman also used DICTION to explore whether good and poor performers differed significantly on the five verbal tone variables. If the theoretical issues being investigated are compatible with DICTION's hard-coded dictionary, the software can be a powerful tool in marketing strategy studies. Like ATLAS.ti, DICTION requires ASCII format files as input.

TextGrab/TextQuest

The growing importance of HTML source files for marketing strategy research, illustrated by the frequency with which Web-based sources are mentioned in Table 1, requires software tools that offer text acquisition capabilities specific to Internet content. Other packages may be moving in that direction, but TextGrab, a software module complementing the content analysis software TextQuest, is specifically designed for the content analysis of websites (Klein 2002), and hence is the pack- age of choice for research primarily interested in this medium. TextGrab is designed to facilitate the acquisition of website data by copying text files (e.g., *.txt *.htm, *.php, *.xml, and *.cfm files) to the user's hard disk. In addition, TextGrab also prepares the files for use with TextQuest by removing HTML tags and translating special characters, such as Greek letters or letters with diacritical marks. The major advantage of TextGrab is that it reduces the time and effort required for using Web-based text data. Disadvantages include the DOS-based nature of TextGrab and the inability of the module to read non-HTML website content, such as JavaScript. A recent review and tutorial of this software combination is available in Garson (2003).

TextQuest is the content-analytic software partner for TextGrab. Like DICTION, TextQuest examines the linguistic style used in the text, but it offers an expanded set of analysis variables (Klein 2002). Text readability, for example, is

assessed by formulas that take sentence and word length (that is, variables explaining most variance of reading behavior) into account. The software currently comprises 60 formulas for 7 languages (31 formulas available for English texts), and numerical results provided by the program include an index value of readability (0-100), reading grade, and reading age classifications. In addition to readability and style analyses, TextQuest provides computer-aided content analysis features, such as search patterns for various category systems (e.g., words, partial words, word sequences, word root chains), concordance, and text indexing. In order to facilitate statistical analysis, TextQuest also provides both automated and interactive coding options, with interactive coding being able to handle ambiguous or negated search patterns. Coding labels and log files are constructed and set up for direct import into statistical packages such as SAS and SPSS. Together, TextQuest and TextGrab can be useful tools for researchers who want to analyze competitor and consumer websites to gauge consumer evaluations, positioning strategies, and other information relevant to strategic marketing decisions.

PolyAnalyst/TextAnalyst

Megaputer markets TextAnalyst as part of its PolyAnalyst program suite. While ATLAS.ti, DICTION, and TextQuest offer category- and dictionary-based coding capabilities, PolyAnalyst/TextAnalyst is a semantic network analysis package that offers automated semantic analysis, summarization, and navigation of unstructured natural-language texts. In addition, TextAnalyst also performs clustering of documents within a text databank, semantic information retrieval, and text exploration around specified subjects. As a suite, PolyAnalyst/TextAnalyst offer comprehensive and flexible qualitative and quantitative functions.

TextAnalyst handles text data in stages. In the preprocessing stage, it strips all supplementary or common words, such as "a," "and," and "the," from the text. It then identifies the stems of words (the base part of words to which endings are added), because analysis of the stems alone provides a clearer picture of each word's frequency and relationship to other words within the analyzed text. While performing the analysis, TextAnalyst also holds information about the complete words and linkages to the source text in case the researchers require access to the source text for in-context assessments of meaning. Once preprocessing is complete, TextAnalyst measures each word's occurrence frequency (excluding common and supplementary words) and the frequency of joint occurrences among semantic elements and generates a tree-like structural representation of the analyzed text. TextAnalyst also calculates statistical weights for each word and their relationships to one another. The resulting semantic network is a concise representation of the analyzed text that can be queried using natural language. For example, asking the question "What is the best minivan?" of articles from 1992 consumer publications triggers a search of the database for the strongest relationships between minivan brands and models and descriptive terms such as "successful." Results from the query are displayed as both a semantic network (tree structure) and as excerpts from the analyzed text, with both providing hyperlinks to the raw text for further analysis.

As with any other software, TextAnalyst has some drawbacks. Similar to ATLAS.ti and DICTION, it requires that all text be in either ASCII or Rich Text Format (RTF). In addition, TextAnalyst requires familiarity with the text data being analyzed in order to make sense of the results, because its analytic procedures focus on the structure of the text instead of its meaning. In its semantic networks, the structure-based process often highlights spurious concepts and relationships that must be eliminated from further consideration through researcher judgment calls. However, if researchers have chosen the text being analyzed carefully and are familiar with the knowledge domain, they can sift through the output and extract insights from the semantic network quickly. The need for human expertise notwithstanding, TextAnalyst can be a valuable tool for researchers who must process large amounts of text data. Its automated processing of text is remarkable in its speed and accuracy.

Future Directions for Text Analysis in Marketing Strategy Research

As indicated by our discussion of text analysis tools, the last few years have witnessed substantial advances in text analysis and automated coding capabilities, and applications have increased in both academic and industry research. One recent example of the increased sensitivity of computer-assisted coding algorithms is the study by Koppel, Argamon, and Shimon (2002), which describes an automated text categorization technique that can infer the unstated gender of the author from formal written documents with approximately 80% accuracy and that can determine if a document is fiction or nonfiction with approximately 98% accuracy. For some types of marketing research, inferring author gender without having to ask can add another explanatory or predictor variable. Consider, for example, research on new automotive models that relies on postings to an auto enthusiast website as text data. Many users of chat rooms and bulletin boards are hesitant to reveal their gender online. Using the approach described by Koppel et al. in conjunction with TextAnalyst, however, it would be feasible to differentiate between male and female consumer postings and gain insight into gender differences in consumer responses to automotive models. Other Internet-based data can be acquired and

analyzed with the capabilities of TextGrab/TextQuest, and the end result can be powerful strategy-shaping insights gleaned solely from text data.

Companies are also using text analysis to address questions that traditional data mining efforts cannot answer. For example, managers at the Fireman's Fund Insurance Company were not able to pinpoint the reasons for rising homeowner claims and suspicious auto claims in traditionally structured quantitative data. By turning to unstructured text files (e.g., claims adjuster notes) and coding the texts to produce predictive models, however, managers found significant explanatory variables of the phenomenon under scrutiny (Ellingsworth and Sullivan 2003). Similarly, coding schemes that capture the frequency and patterns of words and phrases can help managers explore the reasons behind recurring problems of unknown origins. For example, customer service centers that document most of the logged interactions with customers in detail could make fruitful use of text analysis to uncover strategic strengths or weaknesses in a company's current market approach or marketing-variable management.

As these examples indicate, marketing strategy research can benefit greatly from including text analysis as one of its main research methods. Potential future research could assess firm and product performance by analyzing different sources of market conversation. Given that on a product category level, Rosa et al. (1999) found that market stories were able to indicate acceptance and dominant design earlier than sales data, the same might hold true for indicators of firm performance. While profit and market value (stock prices) provide post hoc market assessments of firm performance, market stories might deliver earlier indications of competitive standings and trends. And whereas performance assessments are of interest, the reasons why some companies perform better than others is of crucial importance. One possible avenue for future research, therefore, is to study companies' innovation strategies in published company data in conjunction with marketplace reactions to those strategies, with the goal of identifying the success of a particular strategy for a particular firm in a specific market before making heavy investments in product development projects.

Finally, a long-term goal in adopting text analysis into marketing strategy research should be to create dictionaries and coding categories specific to different marketing strategy research contexts, which could be made available for sharing between research teams and across research studies. By doing so, marketing strategy researchers could enable the comparing of results across studies and allow for increased reliability of coding categories and procedures. For example, just as DICTION incorporates an extensive word list for measuring quantitatively the verbal tone of text passages, an automated coding scheme could assess the strategic orientation (Gatignon and Xuereb 1997) of a particular firm based on its corporate communication materials. Since strategic orientation has been linked to firm performance via innovation and other marketing strategy variables, assessment of strategic orientation could be used in studies to supplement explanatory models of various marketing strategy phenomena. Naturally, such an undertaking would take time and effort, but the potential payoff from developing standardized and reliable coding dictionaries capable of assessing general strategic variables is promising.

José Antonio Rosa is Assistant Professor of Marketing, Weatherhead School of Management, Case Western Reserve University.
Jelena Spanjol is Assistant Professor of Marketing, Mays Business School, Texas A&M University.
Joseph F. Porac is Professor of Management, Stern School of Business, New York University.

Appendix
Selected Websites with Text Analysis Resources

http://www.textanalysis.info/
The website of Harold Klein (creator of the INTEXT and TextQuest software packages). It discusses text analysis and provides a listing of available content analysis software, links to conferences, coding dictionaries, content analysis mailing-list information, and other areas of interest.

http://www.car.ua.edu/
The website of Content Analysis Resources. Information is compiled and maintained by William Evans (University of Alabama). The website contains links to other sites with material on content analysis (CA), references to CA projects and researchers, a listing of available software, CA mailing-list information, and other information.

http://writing.colostate.edu/references/research/content/
This is a CA website hosted at Colorado State University. The site provides an introduction to CA, descriptions of conceptual versus relational CA, a discussion on reliability and validity issues, and an annotated bibliography of CA method pieces and applications.

Notes

1. See Gray and Densten (1998) for a more detailed description of how to integrate quantitative and qualitative content analysis.
2. One hundred sixteen articles containing 5,889 lines of text were sourced from four publications spanning six years (1982-88).
3. *Cycle World* was chosen because it has the largest circulation of any motorcycle industry publication in the United States and because it covers practically all on-road and off-road motorcycle categories sold. For the motorcycle industry, no suitable producer-focused publication was found to portray stories and sensemaking among producers.
4. A related analysis of the local-area network (LAN) market that involved a large number of article abstracts pulled from the ABI-INFORM database was conducted by Theoharakis and Wong (2002) using limited resources and basic coding categories, suggesting that text analysis of emerging product markets is within reach of even small firms or academics working on smaller-scale research projects.

References

Alexa, Melina, and Cornelia Züll (2000), "Text Analysis Software: Commonalities, Differences and Limitations: The Results of a Review." *Quality and Quantity* 34 (3), 299–321.

American Marketing Association (2003), "Honomichl Top 50." *Marketing News* (June 9) H1–48.

Berner, Robert (2003), "No Longer Big Brand on Campus." *Business Week* (September 29), 90.

Bijker, Wiebe E. (1995), *Of Bicycles, Bakelites, and Bulbs: Toward a Theory of Sociotechnical Change.* Cambridge, Mass.: MIT Press.

Denning, Stephen (2001), *The Springboard: How Storytelling Ignites Action in Knowledge-Era Organizations.* Woburn, Mass.: Butterworth-Heinemann.

Ellingsworth, Marty, and Dan Sullivan (2003), "Text Mining Improves Business Intelligence and Predictive Modeling in Insurance." *DM Review* (July). Retrieved July 1, 2004 from http://www.dmreview.com/article_sub.cfm?articleID=6995.

Gabriel, Yiannis (1998), "Stories and Sense-Making." In *Discourse and Organization*, eds. David Grant, Tom Keenoy, and Cliff Oswick, 84–103. London, U.K.: Sage Publications.

Gardner, Howard (1995), *Leading Minds: An Anatomy of Leadership.* New York, N.Y.: Basic Books.

Garson, David G. (2003), "Doing Web-Based Content Profile Analysis: A Tutorial Review of TextGrab and TextQuest." *Social Science Computer Review* 21 (2), 250–60.

Garud, Raghu, and Michael A. Rappa (1994), "A Socio-Cognitive Model of Technology Evolution: The Case of Cochlear Implants." *Organization Science* 5 (August), 344–62.

Gatignon, Hubert, and Jean-Marc Xuereb (1997), "Strategic Orientation of the Firm and New Product Performance." *Journal of Marketing Research* 34 (1) (February), 77–90.

Gladwell, Malcolm (2000), *The Tipping Point: How Little Things Can Make a Big Difference.* Boston, Mass.: Little, Brown and Co.

Gray, Judy H., and Iain L. Densten (1998), "Integrating Quantitative and Qualitative Analysis Using Latent and Manifest Variables." *Quality and Quantity* 32 (4), 419–31.

Hart, Roderick P. (2000), *DICTION 5.0: The Text Analysis Program.* Thousand Oaks, Calif.: Scolari/Sage Publications.

Klein, Harald (2002), *TextGrab/TextQuest.* Rudolstadt, Germany: Social Science Consulting.

Koppel, Moshe, Shlomo Argamon, and Anat Rachel Shimon (2002), "Automatically Categorizing Written Texts by Author Gender." *Literary and Linguistic Computing* 17 (4), 401–12.

Levine, Rick (2000), *The Cluetrain Manifesto: The End of Business as Usual.* Cambridge, Mass.: Perseus Books.

Muhr, Thomas (1997), *ATLAS.ti: Visual Qualitative Data Analysis – Management – Model Building.* Berlin, Germany: Scientific Software Development.

Neuendorf, Kimberly A. (2002), *The Content Analysis Guidebook.* Thousand Oaks, Calif.: Sage Publications.

Ober, Scot, Jensen J. Zhao, Rod Davis, and Melody W. Alexander (1999), "Telling It Like It Is: The Use of Certainty in Public Business Discourse." *Journal of Business Communication* 36 (3), 280–300.

Petroski, Henry (1993), *From Pins to Paper Clips: The Evolution of Useful Things.* New York, N.Y.: Alfred A. Knopf.

Pinch, Trevor, and Wiebe E. Bijker (1987), "The Social Construction of Facts and Artifacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other." In *The Social Construction of Technological Systems,* eds. Wiebe Bijker, Thomas P. Hughes, and Trevor Pinch, 17–50. Cambridge, Mass.: MIT Press.

Porac, Joseph F., Howard Thomas, Fiona Wilson, Douglas Paton, and Alaina Kanfer (1995), "Rivalry and the Industry Model of Scottish Knitwear Producers." *Administrative Science Quarterly* 40 (June), 203–27.

Roberts, Carl W. (2000), "A Conceptual Framework for Quantitative Text Analysis." *Quality and Quantity* 34 (3), 259–74.

Rosa, José Antonio, Kimberly M. Judson, and Joseph F. Porac (2005), "On the Sociocognitive Dynamics between Categories and Product Models in Mature Markets." *Journal of Business Research,* forthcoming.

———, Joseph F. Porac, Jelena Runser-Spanjol, and Michael S. Saxon (1999), "Sociocognitive Dynamics in a Product Market." *Journal of Marketing* 63 (Special Issue), 64–77.

Shaw, Gordon, Robert Brown, and Philip Bromiley (1998), "Strategic Stories: How 3M Is Rewriting Business Planning." *Harvard Business Review* 76 (3), 41–50.

Sirsi, Ajay K., James C. Ward, and Peter H. Reingen (1996), "Microcultural Analysis of Variation in Sharing of Causal Reasoning about Behavior." *Journal of Consumer Research* 22 (March), 345–72.

Sullivan, Dan (2001), *Document Warehousing and Text Mining: Techniques for Improving Business Operations, Marketing, and Sales.* New York, N.Y.: John Wiley & Sons.

Sydserff, Robin, and Pauline Weetman (2002), "Developments in Content Analysis: A Transitivity Index and DICTION Scores." *Accounting, Auditing and Accountability Journal,* 15 (4), 523–45.

Theoharakis, Vasilis, and Veronica Wong (2002), "Marking High-Technology Market Evolution Through the Foci of Market Stories: The Case of Local Area Networks." *Journal of Product Innovation Management* 19 (6), 400–11.

Vaara, Eero (2002), "On the Discursive Construction of Success/Failure in Narratives of Post-merger Integration." *Organization Studies* 23 (2), 211–48.

Venette, Steven J., Timothy L. Sellnow, Patricia A. Lang (2003), "Metanarration's Role in Restructuring Perceptions of Crisis: NHTSA's Failure in the Ford-Firestone Crisis." *Journal of Business Communication* 40 (3), 219–36.

Weick, Karl E. (1995), *Sensemaking in Organizations.* Thousand Oaks, Calif.: Sage Publications.

White, Harrison C. (1981), "Where Do Markets Come From?" *American Journal of Sociology* 87 (November), 517–47.

Züll, Cornelia, and Juliane Landmann (2002), *Computerunterstützte Inhaltsanalyse: Literaturbericht zu neueren Anwendungen.* [*Computer-assisted Content Analysis: Literature Report on More Recent Applications*] Mannheim, Germany: ZUMA.

———, and Peter Ph. Mohler (2001), *Computerunterstützte Inhaltsanalyse: Codierung und Analyse von Antworten auf offene Fragen.* [*Computer-assisted Content Analysis: Coding and Analysis of Answers to Open-ended Questions*] Mannheim, Germany: ZUMA.