Article

AMERICAN MARKETING ASSOCIATION

Journal of Marketing
1-25

DOI: 10.1177/0022242919873106
journals.sagepub.com/home/jmx
SAGE

Uniting the Tribes: Using Text for Marketing Insight

Jonah Berger, Ashlee Humphreys, Stephan Ludwig, Wendy W. Moe, Oded Netzer, and David A. Schweidel

Abstract

Words are part of almost every marketplace interaction. Online reviews, customer service calls, press releases, marketing communications, and other interactions create a wealth of textual data. But how can marketers best use such data? This article provides an overview of automated textual analysis and details how it can be used to generate marketing insights. The authors discuss how text reflects qualities of the text producer (and the context in which the text was produced) and impacts the audience or text recipient. Next, they discuss how text can be a powerful tool both for prediction and for understanding (i.e., insights). Then, the authors overview methodologies and metrics used in text analysis, providing a set of guidelines and procedures. Finally, they further highlight some common metrics and challenges and discuss how researchers can address issues of internal and external validity. They conclude with a discussion of potential areas for future work. Along the way, the authors note how textual analysis can unite the tribes of marketing. While most marketing problems are interdisciplinary, the field is often fragmented. By involving skills and ideas from each of the subareas of marketing, text analysis has the potential to help unite the field with a common set of tools and approaches.

Keywords

computational linguistics, machine learning, marketing insight, interdisciplinary, natural language processing, text analysis, text mining

Online supplement: https://doi.org/10.1177/0022242919873106

The digitization of information has made a wealth of textual data readily available. Consumers write online reviews, answer open-ended survey questions, and call customer service representatives (the content of which can be transcribed). Firms write ads, email frequently, publish annual reports, and issue press releases. Newspapers contain articles, movies have scripts, and songs have lyrics. By some estimates, 80%–95% of all business data is unstructured, and most of that unstructured data is text (Gandomi and Haider 2015).

Such data has the potential to shed light on consumer, firm, and market behavior, as well as society more generally. But, by itself, all this data is just that—data. For data to be useful, researchers must be able to extract underlying insight—to measure, track, understand, and interpret the causes and consequences of marketplace behavior.

This is where the value of automated textual analysis comes in. Automated textual analysis[1] is a computer-assisted methodology that allows researchers to rid themselves of measurement straitjackets, such as scales and scripted questions, and to quantify the information contained in textual data as it naturally occurs. Given these benefits, the question is no longer whether to use automated text analysis but *how* these tools can best be used to answer a range of interesting questions.

This article provides an overview of the use of automated text analysis for marketing insight. Methodologically, text analysis approaches can describe "what" is being said and "how" it is said, using both qualitative and quantitative inquiries with various degrees of human involvement. These

[1] Computer-aided approaches to text analysis in marketing research are generally interchangeably referred to as computer-aided text analysis (Pollach 2012), text mining (Netzer et al. 2012), automated text analysis (Humphreys and Wang 2017), or computer-aided content analysis (Dowling and Kabanoff 1996).

Jonah Berger is Associate Professor of Marketing, Wharton School, University of Pennsylvania, USA (email: jberger@wharton.upenn.edu). Ashlee Humphreys is Associate Professor, Medill School of Journalism, Media, and Integrated Marketing Communications, Northwestern University, USA (email: a-humphreys@northwestern.edu). Stephan Ludwig is Associate Professor of Marketing, University of Melbourne, Australia (email: stephan.ludwig@unimelb.edu.au). Wendy W. Moe is Associate Dean of Master's Programs, Dean's Professor of Marketing, and Co-Director of the Smith Analytics Consortium, University of Maryland, USA (email: wmoe@rhsmith.umd.edu). Oded Netzer is Professor of Business, Columbia Business School, Columbia University, USA (email: onetzer@gsb.columbia.edu). David A. Schweidel is Professor of Marketing, Goizueta Business School, Emory University, USA (email: dschweidel@emory.edu).

approaches consider individual words and expressions, their linguistic relationships within a document (within-text interdependencies) and across documents (across-text interdependencies), and the more general topics discussed in the text. Techniques range from computerized word counting and applying dictionaries to supervised or automated machine learning that helps deduce psychometric and substantive properties of text.

Within this emerging domain, we aim to make four main contributions. First, we illustrate how text data can be used for both prediction and understanding, to gain insight into who produced that text, as well as how that text may impact the people and organizations that consume it. Second, we provide a how-to guide for those new to text analysis, detailing the main tools, pitfalls, and challenges that researchers may encounter. Third, we offer a set of expansive research propositions pertaining to using text as a means to understand meaning making in markets with a focus on how customers, firms, and societies construe or comprehend marketplace interactions, relationships, and themselves. Whereas previous treatments of text analysis have looked specifically at consumer text (Humphreys and Wang 2017), social media communication (Kern et al. 2016), or psychological processes (Tausczik and Pennebaker 2010), we aim to provide a framework for incorporating text into marketing research at the individual, firm, market, and societal levels. By necessity, our approach includes a wide-ranging set of textual data sources (e.g., user-generated content, annual reports, cultural artifacts, government text).

Fourth, and most importantly, we discuss how text analysis can help "unite the tribes." As a field, part of marketing's value is its interdisciplinary nature. Unlike core disciplines such as psychology, sociology, or economics, the marketing discipline is a big tent that allows researchers from different traditions and research philosophies (e.g., quantitative modeling, consumer behavior, strategy, consumer culture theory) to come together to study related questions (Moorman et al. 2019a, b). In reality, however, the field often seems fragmented. Rather than different rowers all simultaneously pulling together, it often feels more like separate tribes, each independently going off in separate directions. Although everyone is theoretically working toward similar goals, there tends to be more communication within groups than between them. Different groups often speak different "languages" (e.g., psychology, sociology, anthropology, statistics, economics, organizational behavior) and use different tools, making it increasingly difficult to have a common conversation. However, text analysis can unite the tribes. Not only does it involve skills and ideas from each of these areas, doing it well *requires* such integration because it borrows ideas, concepts, approaches, and methods from each tribe and incorporates them to achieve insight. In so doing, the approach also adds value to each of the tribes in ways that might not otherwise be possible.

We start by discussing two distinctions that are useful when thinking about how text can be used: (1) whether text reflects or impacts (i.e., says something about the producer or has a downstream impact on something else) and (2) whether text is used for prediction or understanding (i.e., predicting something or understanding what caused something). Next, we explain how text may be used to unite the tribes of marketing. Then we provide an overview of text analysis tools and methodology and discuss key questions and measures of validity. Finally, we close with a future research agenda.

The Universe of Text

Communication is an integral part of marketing. Not only do firms communicate with customers, but customers communicate with firms and one another. Moreover, firms communicate with investors and society communicates ideas and values to the public (through newspapers and movies). These communications generate text or can be transcribed into text.

A simple way to organize the world of textual data is to think about producers and receivers—the person or organization that creates the text and the person or organization who consumes the text (Table 1). While there are certainly other parties that could be listed, some of the main producers and receivers are consumers, firms, investors, and society at large. Consumers write online reviews that are read by other consumers, firms create annual reports that are read by investors, and cultural producers represent societal meanings through the creation of books, movies, and other digital or physical artifacts that are consumed by individuals or organizations.

Consistent with this distinction between text producer and text receiver, researchers may choose to study how text reflects or impacts. Specifically, text reflects information about, and thus can be used to gain insight into, the text *producer* or one can study how text impacts the text *receiver*.

Text as a Reflection of the Producer

Text reflects and indicates something about the text producer (i.e., the person, organization, or context that created it). Customers, firms, and organizations use language to express themselves or achieve desired goals, and as a result, text signals information about the actors, organization, or society that created it and the contexts in which it was created. Like an anthropologist piecing together pottery shards to learn about a distant civilization, text provides a window into its producers.

Take, for example, a social media post in which someone talks about what they did that weekend. The text that person produces provides insight into several facets. First, it provides insight into the individual themselves. Are they introverted or extraverted? Neurotic or conscientious? It sheds light on who they are in general (i.e., stable traits or customer segments; Moon and Kamakura 2017) as well as how they may be feeling or what they may be thinking at the moment (i.e., states). In a sense, language can be viewed as a fingerprint or signature (Pennebaker 2011). Just like brush strokes or painting style can be used to determine who painted a particular painting, researchers use words and linguistic style to infer whether a play was written by Shakespeare, or if a person is depressed (Rude, Gortner, and Pennebaker 2004) or being deceitful

Table 1. Text Producers and Receivers.

Text Producers	Text Receivers: Consumers	Firms	Investors	Institutions/Society
Consumers	• Online reviews (Anderson and Simester 2014; Chen and Lurie 2013; Fazio and Rockledge 2015[a]; Kronrod and Danziger 2013[a]; Lee and Bradlow 2011; Liu, Lee, and Srinivasan 2019[a]; Melumad, Inman, and Pham 2019; Moon and Kamakura 2017; Puranam, Narayan, and Kadiyali 2017) • Social media (Hamilton, Schlosser, and Chen 2017[a]; Netzer et al. 2012; Villarroel Ordenes et al. 2017) • Offline word of mouth (Berger and Schwartz 2011[a]; Mehl and Pennebaker 2003[a])	• Forms and applications (Netzer, Lemaire, and Herzenstein 2019) • Idea-generation contexts (Bayus 2013[a]; Toubia and Netzer 2017) • Social media/brand communities (Herhausen et al. 2019) • Consumer complaints (Ma, Baohung, and Kekre 2015) • Customer language on service calls • Tweeting at companies (Liu, Singh, and Srinivasan 2016[a])	• Stock market reactions to consumer text (Bollen, Mao, and Zeng 2011; Tirunillai and Tellis 2012) • Protests • Petitions	• Crowdsourcing knowledge • Letters to the editor • Online comments section • Activism (e.g., organizing political movements and marches)
Firms	• Owned media (e.g., company website and social media; Villarroel Ordenes et al. 2018) • Advertisements (Fossen and Schweidel 2017[a], 2019; Liaukonyte, Teixeira, and Wilbur 2015[a]; Rosa et al. 1999; Stewart and Furse 1986) • Customer service agents (Packard and Berger 2019; Packard, Moore, and McFerran 2018) • Packaging, including labels • Text used in instructions	• Trade publications (Weber, Heinze, and DeSoucey 2008[a]) • Interfirm communication emails (Ludwig et al. 2016) • White papers	• Financial reports (Loughran and McDonald 2016) • Corporate communications (Hobson, Mayhew, and Venkatachalam 2012) • Chief executive officer letters to shareholders (Yadav, Prabhu, and Chandy 2007	• Editorials by firm stakeholders • Interviews with business leaders
Investors		• Letters to shareholders (Yadav, Prabhu, and Chandy 2007) • Shareholder feedback (Wies et al. 2019)	• Sector reports	
Institutions/ society	• News content (Berger, Kim, and Meyer 2019; Berger and Milkman 2012; Humphreys 2010) • Movies (Berger, Moe, and Schweidel 2019; Eliashberg, Hui, and Zhang 2007, 2014; Toubia et al. 2019) • Songs (Berger and Packard 2018; Packard and Berger 2019) • Books (Akpinar and Berger 2015; Sorescu et al. 2018[a])	• Business section • Specialty magazines (e.g., *Wired*, *Harvard Business Review*)	• *Wall Street Journal* • *Fortune* • Various forms of investment advice that come from media	• Government documents, hearings, and memoranda (Chappell et al. 1997[a]) • Forms of public dialogue or debate

[a]Reference appears in the Web Appendix.

(Ludwig et al. 2016). The same is true for groups, organizations, or institutions. Language reflects something about who they are and thus provides insight into what they might do in the future.

Second, text can provide insight into a person's attitudes toward or relationships with other attitude objects—whether that person liked a movie or hated a hotel stay, for example, or whether they are friends or enemies with someone. Language used in loan applications provides insight into whether people will default (Netzer, Lemaire, and Herzenstein 2019), language used in reviews can provide insight into whether they are fake (Anderson and Simester 2014; Hancock et al. 2007;

Ott, Cardie, and Hancock 2012), and language used by political candidates could be used to study how they might govern in the future.

These same approaches can also be used to understand leaders, organizations, or cultural elites through the text they produce. For example, the words a leader uses reflect who they are as an individual, their leadership style, and their attitudes toward various stakeholders. The language used in ads, on websites, or by customer service agents reflects information about the company those pieces of text represent. Aspects such as brand personality (Opoku, Abratt, and Pitt 2006), how much a firm is thinking about its customers (Packard and Berger 2019), or managers' orientation toward end users (Molner, Prabhu, and Yadav 2019) can be understood through text. Annual reports provide insight into how well a firm is likely to perform in the future (Loughran and McDonald 2016).

Yet beyond single individuals or organizations, text can also be aggregated across creators to study larger social groups or institutions. Given that texts reflect information about the people or organizations that created them, grouping people or organizations together on the basis of shared characteristics can provide insight into the nature of such groups and differences between them. Analyzing blog posts, for example, can shed light on how older and younger people view happiness differently (e.g., as excitement vs. peacefulness; Mogilner, Kamvar, and Aaker 2011). In a comparison of newspaper articles and press releases about different business sectors, text can be used to understand the creation and spread of globalization discourse (Fiss and Hirsch 2005). Customers' language use further gives insight into the consumer sentiment in online brand communities (Homburg, Ehm, and Artz 2015).

More broadly, because texts are shaped by the contexts (e.g., devices, cultures, time periods) in which they were produced, they also reflect information about these contexts. In the case of culture, U.S. culture values high-arousal positive affective states more than East Asian cultures (Tsai 2007), and these differences may show up in the language these different groups use. Similarly, whereas members of individualist cultures tend to use first-person pronouns (e.g., "I"), members of collectivist cultures tend to use a greater proportion of third-person pronouns (e.g., "we").

Across time, researchers were able to examine whether the national mood changed after the September 11 attacks by studying linguistic markers of psychological change in online diaries (Cohn, Mehl, and Pennebaker 2004). The language used in news articles, songs, and public discourse reflects societal attitudes and norms, and thus analyzing changes over time can provide insight into aspects such as attitudes toward women and minorities (Boghrati and Berger 2019; Garg et al. 2018) or certain industries (Humphreys 2010). Journal articles provide a window into the evolution of topics within academia (Hill and Carley 1999). Books and movies serve as similar cultural barometers and could be used to shed light on everything from cultural differences in customs to changes in values over time.

Consequently, text analysis can provide insights that may not be easily (or cost-effectively) obtainable through other methods. Companies and organizations can use social listening (e.g., online reviews and blog posts) to understand whether consumers like a new product, how customers feel about their brand, what attributes are relevant for decision making, or what other brands fall in the same consideration set (Lee and Bradlow 2011; Netzer et al. 2012). Regulatory agencies can determine adverse reactions to pharmaceutical drugs (Feldman et al. 2015; Netzer et al. 2012), public health officials can gauge how bad the flu will be this year and where it will hit the hardest (Alessa and Faezipour 2018), and investors can try to predict the performance of the stock market (Bollen, Mao, and Zeng 2011; Tirunillai and Tellis 2012).

Text's Impact on Receivers

In addition to reflecting information about the people, organizations, or society that created it, text also impacts or shapes the attitudes, behavior, and choices of the audience that consumes it. For example, take the language used by a customer service agent. While that language certainly reflects something about that agent (e.g., their personality, how they are feeling that day), how they feel toward the customer, and what type of brand they represent, that language also impacts the customer who receives it (Packard and Berger 2019; Packard, Moore, and McFerran 2018). It can change customer attitudes toward the brand, influence future purchase, or affect whether customers talk about the interaction with their friends. In that sense, language has a meaningful and measurable impact on the world. It has consequences.

This can be seen in a myriad of different contexts. Ad copy shapes customers' purchase behavior (Stewart and Furse 1986), newspaper language changes customers' attitudes (Humphreys and LaTour 2013), trade publications and consumer magazines shift product category perceptions (e.g., Rosa et al. 1999), movie scripts shape audience reactions (Berger, Kim, and Meyer 2019; Eliashberg, Hui, and Zhang 2014; Reagan et al. 2016), and song lyrics shape song market success (Berger and Packard 2018; Packard and Berger 2019). The language used in political debates shapes which topics get attention (Berman et al. 2019), the language used in conversation shapes interpersonal attitudes (Huang et al. 2017), and the language used in news articles shapes whether people read (Berger, Moe, and Schweidel, 2019b) or share (Berger and Milkman 2012) them.

Firms' language choice has impact as well. For example, nuances in language choices by firms when responding to customer criticism online directly impacts consumers and, thus, the firms' success in containing social media firestorms (Herhausen et al. 2019). Language used in YouTube ads is correlated with their virality (Tellis et al. 2019). Shareholder complaints about nonfinancial concerns and topics that receive high media attention substantially increase firms' advertising investments (Wies et al. 2019).

Note that while the distinction between text reflecting and impacting is a useful one, it is not an either/or. Text almost always simultaneously reflects *and* impacts. Text always reflects information about the actor or actors that created it, and as long as some audience consumes that text, it also impacts that audience.

Despite this relationship, researchers studying reflection versus impact tend to use text differently. Research that examines what the text reflects often treats it as a dependent variable and investigates how it relates to the text creator's personality, the social groups they belong to, or the time period or culture in which it was created.

Research that examines how text impacts others often treats it as an independent variable, examining if and how text shapes outcomes such as purchase, sharing, or engagement. In this framework, textual elements are linked with outcomes that are believed to be theoretical consequences of the textual components or some latent variable that they are thought to represent.

Contextual Influences on Text

Importantly, text is also shaped by contextual factors; thus, to better understand its meaning and impact, it is important to understand the broader situation in which it was produced. Context can affect content in three ways: through technical constraints and social norms of the genre, through shared knowledge specific to the speaker and receiver, and through prior history.

First, different types of texts are influenced by formal and informal rules and norms that shape the content and expectations about the message. For example, newspaper genres such as opinion pieces or feature stories will contain a less "objective" point of view than traditional reporting (Ljung 2000). Hotel comment cards and other feedback are usually dominated by more extreme opinions. On Snapchat and other social media platforms, messages are relatively recent, short, and often ephemeral. In contrast, online reviews can be longer and are often archived dating back several years. Synchronic text exchanges, in which two individuals interactively communicate in real time may be more informal and contain dialogue of short statements and phatic responses (i.e., communication such as "Hi," which serves a social function) that indicate affiliation rather than semantic content (Kulkarni 2014). Some genres (e.g., social media) are explicitly public, whereas on others, such as blogs, information that is more private may be conveyed.

Text is also shaped by technological constraints (e.g., the ability to like or share) and physical constraints (e.g., character length limitations). Tweets, for example, necessarily have 280 characters or fewer, which may shape the ways in which they are used to communicate. Mobile phones have constraints on typing and may shape the text that people produce on them (Melumad, Inman, and Pham 2019; Ransbotham, Lurie, and Liu 2019).

Second, the relationship between the text producer and consumer may affect what is said (or, more often, unsaid). If the producer and consumer know each other well, text may be relatively informal (Goffman 1959) and lack explicit information that a third party would need to make sense of the conversation (e.g., past events, known likes/dislikes). If both have an understanding of the goal of the communication (e.g., that the speaker wants to persuade the receiver), this may shape the content but be less explicit.

These factors are important to understand when interpreting the content of the text itself. Content has been shown to be shaped by the creator's intended audience (Vosoughi, Roy, and Aral 2018) and anticipated effects on the receiver (Barasch and Berger 2014). Similarly, what consumers share with their best friend may be different (e.g., less impacted by self-presentational motivations) than what they post online for everyone to see.[2] Firms' annual reports may be shaped by the goals of appearing favorably to the market. What people say on a customer service call may be driven by the goal of getting monetary compensation. Consumer protests online are meant to inspire change, not merely inform others.

Finally, history may affect the content of the text. In message boards, prior posts may shape future posts; if someone raised a point in a previous post, the respondent will most likely refer to the point in future posts. If retweets are included in an analysis, this will bias content toward most circulated posts. More broadly, media frames such as #metoo or #blacklivesmatter might make some concepts or facts more accessible to speakers and therefore more likely to emerge in text, even if seemingly unrelated (McCombs and Shaw 1972; Xiong, Cho, and Boatwright 2019).

Using Text for Prediction Versus Understanding

Beyond reflecting information about the text creator and shaping outcomes for the text recipient, another useful distinction is whether text is used for prediction or understanding.

Prediction

Some text research is predominantly interested in prediction. Which customer is most likely to default on their loan (Netzer, Lemaire, and Herzenstein 2019)? Which movie will sell the most tickets (Eliashberg et al. 2014)? How will the stock market perform (Bollen, Mao, and Zeng 2011; Tirunillai and Tellis 2012)? Whether focusing on individual-, firm-, or market-level outcomes, the goal is to predict with the highest degree of accuracy. Such work often takes many textual features and uses machine learning or other methods to combine these features in a way that achieves the best prediction. The authors care less

[2] Note that intermediaries can amplify (e.g., retweet) an original message and may have different motivations than the text producer.

about any individual feature and more about how the set of observable features can be combined to predict an outcome.

The main difficulty involved with using text for predictions is that text can generate hundreds and often thousands of features (words) that are all potential predictors for the outcome of interest. In some cases, the number of predictors is larger than the number of observations, making traditional statistical predictive models largely impractical. To address this issue, researchers often resort to machine learning–type methods, but overfitting needs to be carefully considered. In addition, inference with respect to the role of each word in the prediction can be difficult. Methods such as feature importance weighing can help extract some inference from these predictive models.

Understanding

Other research is predominantly interested in using text for understanding. How does the language consumers use shape word of mouth's impact (Packard and Berger 2017)? Why do some online posts get shared, songs become popular, or brands engender greater loyalty? How do cultural attitudes or business practices change? Whether focusing on individual-, firm-, or market-level outcomes, the goal is to understand why or how something occurred. Such work often involves examining only one or a small number of textual features or aspects that link to underlying psychological or sociological processes and aims to understand which features are driving outcomes and why.

One challenge with using textual data for understanding is drawing causal inferences from observational data. Consequently, work in this area may augment field data with experiments to allow key independent variables to be manipulated. Another challenge is interpreting relationships with textual features (we discuss this further in the closing section). Songs that use more second-person pronouns are more popular (Packard and Berger 2019), for example, but this relationship alone does not necessarily explain *why* this is the case; second-person pronouns may indicate several things. Consequently, deeper theorizing, examination of links observed in prior research, or further empirical work is often needed.

Note that research that can use either a prediction or understanding lens to study either what text reflects or what it impacts. On the prediction side, researchers interested in what text reflects could use it to predict states or traits of the text creator such as customer satisfaction, likelihood of churn, or brand personality. Researchers interested in the impact of text could predict how text will shape outcomes such as reading behavior, sharing, or purchase among consumers of that text.

On the understanding side, someone interested in what text reflects could use it to shed light on why people might use certain personal pronouns when they are depressed or why customers might use certain types of emotional language when they are talking to customer service. Someone interested in the impact of text could use it to understand why text that evokes different emotions might be more likely to be read or shared.

Furthermore, while most research tends to focus on either prediction or understanding, some work integrates both aspects. Netzer, Lemaire, and Herzenstein (2019), for example, both use a range of available textual features to predict whether a given person will default on a loan and analyze the specific language used by people who tend to default (e.g., language used by liars).

Uniting the Tribes of Marketing

Regardless of whether the focus is on text reflection versus impact, or prediction versus understanding, doing text analysis well requires integrating skills, techniques, and substantive knowledge from different areas of marketing. Furthermore, textual analysis opens up a wealth of opportunity for each of these areas as well.

Take consumer behavior. While hypothetical scenarios can be useful, behavioral economics has recently gotten credit for many applications of social or cognitive psychology because these researchers have demonstrated phenomena in the field. Given concerns about replication, researchers have started to look for new tools that enable them to ensure validity and increase relevance to external audiences. Previously, use of secondary data was often limited because it addressed the "what" but not the "why" (i.e., what people bought or did, but not why they did so). But text can provide a window into the underlying process. Online reviews, for example, can be used to understand why someone bought one thing rather than another. Blog posts can help marketers understand consideration sets (Lee and Bradlow 2011; Netzer et al. 2012) and the customer journey (Li and Du 2011). Text even helps address the age-old issue of telling more than we can know (Nisbett and Wilson 1977). While people may not always know why they did something, their language often provides traces of explanation (Pennebaker 2011), even beyond what they can consciously articulate.

This richness is attractive to more than just behavioral researchers. Text opens a large-scale window into the world of "why" in the field and does so in a scalable manner. Quantitative modelers are always looking for new data sources and tools to explain and predict behavior. Unstructured data provides a rich set of predictors that are often readily available, at large scale, and able to be combined with structured measures as either dependent variables or independent variables. Text, through product reviews, user-driven social media activity, and firm-driven marketing efforts, provides data in real time that can shed light on consumer needs/preferences. This offers an alternative or supplement to traditional marketing research tools. In many cases, text can be retraced to an individual, allowing distinction between individual differences and dynamics. It also offers a playground where new methodologies from other disciplines can be applied (e.g., deep learning; LeCun, Bengio, and Hinton 2015; Liu et al. 2019).

Marketing strategy researchers want logic by which business can achieve its marketing objectives and to better understand what affects organizational success. A primary challenge to these researchers is to obtain reliable and generalizable survey or field data about factors that lie deep in the firm's culture and structure or that are housed in the mental models and beliefs of marketing leaders and employees. Text analysis offers an objective and systematic solution to assess constructs in naturally occurring data (e.g., letters to shareholders, press releases, patent text, marketing messages, conference calls with analysts) that may be more valid. Likewise, marketing strategy scholars often struggle with valid measures of a firm's marketing assets, and text may be a useful tool to understand the nature of customer, partner, and employee relationships and the strength of brand sentiments. For example, Kübler, Colicev, and Pauwels (2017) use dictionaries and support vector machine methods to extract sentiment and relate it to consumer mindset metrics.

Scholars who draw from anthropology and sociology have long examined text through qualitative interpretation and content analysis. Consumer culture theory–oriented marketing researchers are primarily interested in understanding underlying meanings, norms, and values of consumers, firms, and markets in the marketplace. Text analysis provides a tool for quantifying qualitative information to measure changes over time or make comparisons between groups. Sociological and anthropological researchers can use automated text analysis to identify important words, locate themes, link them to text segments, and examine common expressions in their context. For example, to understand consumer taste practices, Arsel and Bean (2013) use text analysis to first identify how consumers talk about different taste objects, doings, and meanings in their textual data set (comments on a website/blog) before analyzing the relationship between these elements using interview data.

For marketing practitioners, textual analysis unlocks the value of unstructured data and offers a hybrid between qualitative and quantitative marketing research. Like qualitative research, it is rich, exploratory, and can answer the "why," but like quantitative research, it benefits from scalability, which often permits modeling and statistical testing. Textual analysis enables researchers to explore open-ended questions for which they do not know the range of possible answers a priori. With text, scholars can answer questions that they did not ask or for which they did not know the right outcome measure. Rather than forcing on participants a certain scale or set of outcomes from which to select, for example, marketing researchers can instead ask participants broad questions, such as why they like or dislike something, and then use topic modeling tools such as latent Dirichlet allocation (LDA; explained in detail subsequently) to discover the key underlying themes.

Importantly, while text analysis offers opportunities for a variety of research traditions, such opportunities are more likely to be realized when researchers work across traditional subgroups. That is, the benefits of computer-aided text analysis are best realized if we include both quantitative, positivist analyses of content and qualitative, interpretive analyses of discourse. Quantitative researchers, for example, have the skills to build the right statistical models, but they can benefit from behavioral and qualitative researchers' ability to link words to underlying psychological or social processes as well as marketing strategy researchers' understanding of organizational and marketing activities driving firm performance. This is true across all of the groups.

Thus, to really extract insights from textual data, research teams must have the interpretative skills to understand the meaning of words, the behavioral skills to link them to underlying psychological processes, the quantitative skills to build the right statistical models, and the strategy skills to understand what these findings mean for firm actions and outcomes. We outline some potential areas for fruitful collaboration in "Future Research Agenda" section.

Text Analysis Tools, Methods, and Metrics

Given the recent work using text analysis to derive marketing insight, some researchers may wonder where to start. This section reviews methodologies often used in text-based research. These include techniques needed to convert text into constructs in the research process as well as procedures needed to incorporate extracted textual information into subsequent modeling and analyses. The objective of this section is not to provide a comprehensive tutorial but, rather, to expose the reader to available techniques, discuss when different methods are appropriate, and highlight some of the key considerations in applying each method.

The process of text analysis involves several steps: (1) data preprocessing, (2) performing a text analysis of the resulting data, (3) converting the text into quantifiable measures, and (4) assessing the validity of the extracted text and measures. Each of these steps may vary depending on the research objective. Table 2 provides a summary of the different steps involved in the text analysis process from preprocessing to commonly used tools and measures and validation approaches. Table 2 can serve as a starter kit for those taking their first steps with text analysis.

Data Preprocessing

Text is often unstructured and "messy," so before any formal analyses can take place, researchers must first preprocess the text itself. This step provides structure and consistency so that the text can be used systematically in the scientific process. Common software tools for text analysis include Python (https://www.nltk.org/) and R (https://cran.r-project.org/web/packages/quanteda/quanteda.pdf, https://quanteda.io/). For both software platforms, a set of relatively easy-to-use tools has been developed to perform most of the data preprocessing steps. Some programs, such as Linguistic Inquiry and Word Count (LIWC; Tausczik and Pennebaker 2010) and WordStat (Peladeau 2016), require minimal preprocessing. We detail the data preprocessing steps next (for a summary of the steps, see Table 3).

Table 2. The Text Analysis Workflow.

Data Preprocessing	Common Tools	Measurement	Validity
• ***Data acquisition***: Obtain or download (often in an HTML format) text. • ***Tokenization***: Break text into units (often words and sentences) using delimiters (e.g., periods). • ***Cleaning***: Remove nonmeaningful text (e.g., HTML tags) and nontextual information. • ***Removing stop words***: Eliminate common words such as "a" or "the" that appear in most documents. • ***Spelling***: Correct spelling mistakes using common spellers. • ***Stemming and lemmatization***: Reduce words into their common stem or lemma.	• ***Entity extraction***: Tools used to extract the meaning of one word at a time or simple cooccurrence of words. These tools include dictionaries; part-of-speech classifiers; many sentiment analysis tools; and, for complex entities, machine learning tools. • ***Topic modeling***: Topic modeling can identify the general topics (described as a combination of words) that are discussed in a body of text. Common tools include LDA and PF. • ***Relation extraction***: Going beyond entity extraction, the researcher may be interested in identifying textual relationships among extracted entities. Relation extraction often requires the use of supervised machine learning approaches.	• ***Count measures***: The set of measures used to represent the text as count measures. The tf-idf measure allows the researcher to control for the popularity of the word and the length of the document. • ***Similarity measures***: Cosine similarity and the Jaccard index are often used to measure the similarity of the text between documents. • ***Accuracy measures***: Often used relative to human-coded or externally validated documents. The measures of recall, precision, F1, and the area under the curve of the receiver operating characteristic curve are often used. • ***Readability measures***: Measures such as the simple measure of gobbledygook (SMOG) are used to assess the readability level of the text.	• ***Internal Validity*** – Construct: Dictionary validation and sampling-and-saturation procedures ensure that constructs are correctly operationalized in text. – Concurrent: Compare operationalizations with prior literature. – Convergent: Multiple operationalizations of key constructs. – Causal: Control for factors related to alternative hypotheses. • ***External Validity*** – Predictive: Use conclusions to predict key outcome variable (e.g., sales, stock price). – Generalizability: Replicate effects in other domains. – Robustness: Test conclusions on holdout samples (k-fold); compare different categories within the data set.

Note: PF = Poisson factoring.

Data acquisition. Data acquisition can be well defined if the researcher is provided with a set of documents (e.g., emails, quarterly reports, a data set of product reviews) or more open-ended if the researcher is using a web scraper (e.g., Beautiful Soup) that searches the web for instances of a particular topic or a specific product. When scraping text from public sources, researchers should abide by the legal guidelines for using the data for academic or commercial purposes.

Tokenization. Tokenization is the process of breaking the text into units (often words and sentences). When tokenizing, the researcher needs to determine the delimiters that define a token (space, period, semicolon, etc.). If, for example, a space or a period is used to determine a word, it may produce some nonsensical tokens. For example, "the U.S." may be broken to the tokens "the," "U," and "S." Most text-mining software has smart tokenization procedures to alleviate such common problems, but the researcher should pay close attention to instances that are specific to the textual corpora. For cases that include paragraphs or threads, depending on the research objective, the researcher may wish to tokenize these larger units of text as well.

Cleaning. HTML tags and nontextual information, such as images, are cleaned or removed from the data set. The cleaning needs may depend on the format in which the data was provided/extracted. Data extracted from the web often requires heavier cleaning due to the presence of HTML tags. Depending on the purpose of the analysis, images and other nontextual information may be retained. Contractions such as "isn't" and "can't" need to be expanded at this step. In this step, researchers should also be mindful of and remove phrases automatically generated by computers that may occur within the text (e.g., "html").

Removing stop words. Stop words are common words such as "a" and "the" that appear in most documents but often provide no significant meaning. Common text-mining tools (e.g., the tm, quanteda, tidytext, and tokenizers package in R; the Natural Language Toolkit package in Python; exclusion words in WordStat) have a predefined list of such stop words that can be amended by the researcher. It is advisable to add common words that are specific to the domain (e.g., "Amazon" in a corpora of Amazon reviews) to this list. Depending on the research objective, stop words can sometimes be very meaningful, and researchers may wish to retain them for their analysis. For example, if the researcher is interested in extracting not only the content of the text but also writing style (e.g., Packard, Moore, and McFerran 2018), stop words can be very informative (Pennebaker 2011).

Spelling. Most text-mining packages have prepackaged spellers that can help correct spelling mistakes (e.g., the Enchant speller). In using these spellers, the researcher should be aware of language that is specific to the domain and may not appear in the speller—or even worse, that the speller may incorrectly "fix." Moreover, for some analyses the researcher may want to record the number of spelling mistakes as an additional

Table 3. Data Preprocessing Steps.

Data Processing Step	Issues to Consider	Illustration
Data acquisition	• Is the data readily available in textual format or does the research needs to use a web scraper to find the data? • What are the legal guidelines for using the data (particularly relevant for web-scraped data)?	Tweets mentioning different brands from the same category during a particular time frame are downloaded from Twitter.
Tokenization	• What is the unit of analysis (word, sentence, thread, paragraph)? • Use smart tokenization for delimiters and adjust to specific unique delimiters found in the corpora.	The unit of analysis is the individual tweet. The words in the tweet are the tokens of the document.
Cleaning	• Web-scraped data often requires cleaning of HTML tags and other symbols. • Depending on the research objective, certain textual features (e.g., advertising on the page) may or may not be cleaned. • Expansion of contractions such as "isn't" to "is not."	URLs are removed and emojis/emoticons are converted to words.
Removing stop word	• Use a stop word list available by the text-mining software, but adapt it to a specific application by adding/removing relevant stop words. • If the goal of the analysis is to extract writing style, it is advisable to keep all/some of the stop words.	Common words are removed. The remaining text contains brand names, nouns, verbs, adjectives, and adverbs.
Spelling	• Can use commonly used spellers in text-mining packages (e.g., the Enchant speller). • Language that is specific to the domain may be erroneously coded as a spelling mistake. • May wish to record the number of spelling mistakes as an additional textual measure.	Spelling mistakes are removed, enabling analysis into consumer perceptions (manifest through word choice) of different brands.
Stemming and lemmatization	• Can use commonly used stemmers in text-mining packages (e.g., Porter stemmer). • If the goal of the analysis is to extract writing style, stemming can mask the tense used.	Verbs and nouns are "standardized" by reducing to their stem or lemma.

textual measure reflecting important states or traits of the communicator (e.g., Netzer, Lemaire, and Herzenstein 2019).

Stemming and lemmatization. Stemming is the process of reducing the words into their word stem. Lemmatization is similar to stemming, but it returns the proper lemma as opposed to the word's root, which may not be a meaningful word. For example, with stemming, the entities "car" and "cars" are stemmed to "car," but "automobile" is not. In lemmatization, the words "car," "cars," and "automobile" are all reduced to the lemma "automobile." Several prepackaged stemmers exist in most text-mining tools (e.g., the Porter stemmer). Similar to stop words, if the goal of the analysis is to extract the writing style, one may wish to skip the stemming step, because stemming often masks the tense used.

Text Analysis Extraction

Once the data has been preprocessed, the researcher can start analyzing the data. One can distinguish between the extraction of individual words or phrases (entity extraction), the extraction of themes or topics from the collective set of words or phrases in the text (topic extraction), and the extraction of relationships between words or phrases (relation extraction). Table 4 highlights these three types of analysis, the typical research questions investigated with each approach, and some commonly used tools.

Entity (word) extraction. At the most basic level, text mining has been used in marketing to extract individual entities (i.e., count words) such as person, location, brands, product attributes, emotions, and adjectives. Entity extraction is probably the most commonly used text analysis approach in marketing academia and practice, partly due to its relative simplicity. It allows the researcher to explore both what was written (the content of the words) as well as how it was written (the writing style). Entity extraction can be used (1) to monitor discussions on social media (e.g., numerous commercial companies offer buzz monitoring services and use entity extraction to track how frequently a brand is being mentioned across alternative social media), (2) to generate a rich set of entities (words) to be used in a predictive model (e.g., which words or entities are associated with fake or fraudulent statements), and (3) as input to be used with dictionaries to extract more complex forms of textual expressions, such as a particular concept, sentiment, emotion, or writing style.

In addition to programming languages such as Python and R's tm tool kits, software packages such as WordStat make it possible to extract entities without coding. Entity extraction can also serve as input in commonly used dictionaries or lexicons. Dictionaries (i.e., a predefined list of words, such as a list of brand names) are often used to classify entities into the categories (e.g., concepts, brands, people, categories, locations). In more formal text, capitalization can be used to help

Table 4. Taxonomy of Text Analysis Tools.

Approach	Common Tools	Research Questions	Benefits	Limitations and Complexities	Marketing Examples
Entity (word) extraction: Extracting and identifying a single word/n-gram	• Named entity extraction (NER) tools (e.g., Stanford NER) • Dictionaries and lexicons (e.g., LIWC, EL 2.0, SentiStrength, VADER) • Rule-based classification • Linguistic-based NLP tools • Machine learning classification tools (conditional random fields, hidden Markov models, deep learning)	• Brand buzz monitoring • Predictive models where text is an input • Extracting psychological states and traits • Sentiment analysis • Consumer and market trends • Product recommendations	• Can extract a large number of entities • Can uncover known entities (people, brands, locations) • Can be combined with dictionaries to extract sentiment or linguistic styles • Relatively simple to use	• Can be unwieldy due to the large number of entities extracted • Some entities have multiple meanings that are difficult to extract (e.g., the laundry detergent brand "All") • Slang and abbreviations make entity extraction more difficult in social media • Machine learning tools may require large human-coded training data • Can be limited for sentiment analysis	• Lee and Bradlow (2011) • Berger and Milkman (2012) • Ghose et al. (2012)[a] • Tirunillai and Tellis (2012) • Humphreys and Thompson (2014)[a] • Berger, Moe, and Schweidel (2019) • Packard, Moore, and McFerran (2018)
Topic extraction: Extracting the topic discussed in the text	• LSA • LDA • PF • LDA2vec word embedding	• Summarizing the discussion • Identifying consumer and market trends • Identifying customer needs	• Topics often provide useful summarization of the data • Data reduction permits the use of traditional statistical methods in subsequent analysis • Easy-to-assess dynamics	• The interpretation of the topics can be challenging • No clear guidance on the selection of the number of topics • Can be difficult with short text (e.g., tweets)	• Tirunillai and Tellis (2014) • Büschken and Allenby (2016) • Puranam, Narayan, and Kadiyali (2017) • Berger and Packard (2018) • Liu and Toubia (2018) • Toubia et al. (2019) • Zhong and Schweidel (2019) • Ansari, Li, and Yang (2018)[a] • Timoshenko and Hauser (2019) • Liu, Singh, and Srinivasan (2016)[a] • Liu, Lee, and Srinivasan (2019)[a]
Relation extraction: Extracting and identifying relationships among words	• Co-occurrence of entities • Handwritten rule • Supervised machine learning • Deep learning • Word2vec word embedding • Stanford Sentence and Grammatical Dependency Parser	• Market mapping • Identifying problems mentioned with specific product features • Identifying sentiment for a focal entity • Identifying which product attributes are mentioned positively/negatively • Identifying events and consequences (e.g., crisis) from consumer- or firm-generated text • Managing service relationships	• Relaxes the bag-of-words assumption of most text-mining methods • Relates the text to a particular focal entity • Advances in text-mining methods will offer new opportunities in marketing	• Accuracy of current approaches is limited • Complex relationships may be difficult to extract • It is advised to develop domain-specific sentiment tools as sentiment signals can vary from one domain to another	• Netzer et al. (2012) • Toubia and Netzer (2017) • Boghrati and Berger (2019)

[a]Reference appears in the Web Appendix.

extract known entities such as brands. However, in more casual text, such as social media, such signals are less useful. Common dictionaries include LIWC (Pennebaker et al. 2015), EL 2.0 (Rocklage, Rucker, and Nordgren 2018), Diction 5.0, or General Inquirer for psychological states and traits (for example applications, see Berger and Milkman [2012]; Ludwig et al. [2013]; Netzer, Lemaire, and Herzenstein [2019]).

Sentiment dictionaries such as Hedonometer (Dodds et al. 2011), VADER (Hutto and Gilbert 2014), and LIWC can be used to extract the sentiment of the text. One of the major limitations of the lexical approaches for sentiment analysis commonly used in marketing is that they apply a "bag of words" approach—meaning that word order does not matter—and rely solely on the cooccurrence of a word of interest (e.g., "brand") with positive or negative words (e.g., "great," "bad") in the same textual unit (e.g., a review). While dictionary approaches may be an easy way to measure constructs and comparability across data sets, machine learning approaches trained by human-coded data (e.g., Borah and Tellis 2016; Hartmann et al. 2018; Hennig-Thurau, Wiertz, and Feldhaus 2015) tend to be the most accurate way of measuring such constructs (Hartmann et al. 2019), particularly if the construct is complex or the domain is uncommon. For this reason, researchers should carefully weigh the trade-off between empirical fit and theoretical commensurability, taking care to validate any dictionaries used in the analysis (discussed in the next section).

A specific type of entity extraction includes linguistic-type entities such as part-of-speech tagging, which assigns a linguistic tag (e.g., verb, noun, adjective) to each entity. Most text analysis tools (e.g., the tm package in R, the Natural Language Toolkit package in Python) have a built-in part-of-speech tagging tool. If no predefined dictionary exists, or the dictionary is not sufficient for the extraction needed, one could add handcrafted rules to help define entities. However, the list of rules can become long, and the task of identifying and writing the rules can be tedious. If the entity extraction by dictionaries or rules is difficult or if the entities are less defined, machine learning–supervised classification approaches (e.g., conditional random fields [Netzer et al. 2012], hidden Markov models) or deep learning (Timoshenko and Hauser 2019) can be used to extract entities. The limitation of this approach is that often a relatively large hand-coded training data set needs to be generated.

To allow for a combination of words, entities can be defined as a set of consecutive words, often referred to as n-grams, without attempting to extract the relationship between these entities (e.g., the consecutive words "credit card" can create the unigram entities "credit" and "card" as well as the bigram "credit card"). This can be useful if the researcher is interested in using the text as input for a predictive model.

If the researcher wishes to extract entities while understanding the context in which the entities were mentioned in the text (thus avoiding the limitation of the bag-of-words approach), the emerging set of tools of word2vec or word embedding (Mikolov et al. 2013) can be employed. Word2vec maps each word or entity to a vector of latent dimensions called embedding vector based on the words with which each focal word appears. This approach allows the researcher not only to extract words but also to understand the similarity between words based on the similarities between the embedding vectors (or the similarities between the sentences in which each word appears). Thus, unlike the previous approaches discussed thus far, word2vec preserves the context in which the word appeared. While word embedding statistically captures the context in which a word appears, it does not directly linguistically "understand" the relationships among words.

Topic modeling. Entity extraction has two major limitations: (1) the dimensionality of the problem (often thousands of unique entities are extracted) and (2) the interpretation of many entities. Several topic modeling approaches have been suggested to overcome these limitations. Similar to how factor analysis identifies underlying themes among different survey items, topic modeling can identify the general topics (described as a combination of words) that are discussed in a body of text. This text summarization approach increases understanding of document content and is particularly useful when the objective is insight generation and interpretation rather than prediction (e.g., Berger and Packard 2018; Tirunillai and Tellis 2014). In addition, monitoring topics, as opposed to words, makes it easier to assess how discussion changes over time (e.g., Zhong and Schweidel 2019).

Methodologically, topic modeling mimics the data-generating process in which the writer chooses the topic she wants to write about and then chooses the words to express these topics. Topics are defined as word distributions that commonly co-occur and thus have a certain probability of appearing in a topic. A document is then described as a probabilistic mixture of topics.

The two most commonly used tools for topic modeling are LDA (Blei, Ng, and Jordan 2003) and Poisson factorization (PF; Gopalan, Hofman and Blei 2013). The predominant approach prior to LDA and PF was the support-vector-machine latent semantic analysis (LSA) approach. While LSA is simpler and faster to implement than LDA and PF, it requires larger textual corpora and often achieves lower accuracy levels. Other approaches include building an ontology of topics using a combination of human classification of documents as seeding for a machine learning classification (e.g., Moon and Kamakura 2017). Whereas LDA is often simpler to apply than PF, PF has the advantage of not assuming that the topic probabilities must sum to one. That is, some documents may have more topic presences than others, and a document can have multiple topics with high likelihood of occurrence. In addition, PF tends to be more stable with shorter text. Büschken and Allenby (2016) relax the common bag-of-words assumption underlying the traditional LDA model and leverage the within-sentence dependencies of online reviews. LDA2vec is another approach to assess topics while accounting for the sequence context in which the word appears (Moody 2016). In the context of search queries, Liu and Toubia (2018) further extend the LDA approach to hierarchical LDA for cases in which related documents (queries and search results) are used to extract the topics. Furthermore, the researcher can use an unsupervised or seeded LDA approach to incorporate prior knowledge in the construction and interpretation of the topics (e.g., Puranam, Narayan, and Kadiyali 2017; Toubia et al. 2019).

While topic modeling methods often produce very sensible topics, because topics are selected solely based on a statistical approach, the selection of the number of topics and the interpretation of some topics can be challenging. It is recommended to combine statistical approaches (e.g., the perplexity measure, which is a model fit–based measure) and researcher judgment when selecting the number of topics.

Relation extraction. At the most basic level, relationships between entities can be captured by the mere co-occurrence of entities (e.g., Boghrati and Berger 2019; Netzer et al. 2012; Toubia and Netzer 2017). However, marketing researchers are often more interested in identifying textual relationships among extracted entities, such as the relationships between products, attributes, and sentiments. Such relationships are often more relevant for the firm than merely measuring the volume of brand mentions or even the overall brand sentiment. For example, researchers may want to identify whether consumers mentioned a particular problem with a specific product feature. Feldman et al. (2015) and Netzer et al. (2012) provide such examples by identifying the textual relationships between drugs and adverse drug reactions that imply that a certain drug may cause a particular adverse reaction.

Relation extraction also offers a more advanced route to capture sentiment by providing the link between an entity of interest (e.g., a brand) and the sentiment expressed, beyond their mere cooccurrence. Relation extraction based on the bag-of-words approach, which treats the sentence as a bag of unsorted words and searches for word cooccurrence, is limited because the cooccurrence of words may not imply a relationship. For example, the cooccurrence of a drug (e.g., Advil) with a symptom (e.g., headache) may refer to the symptom as a side effect of the drug or as the effect the drug is aiming to alleviate. Addressing such relationships requires identifying the sequence of words and the linguistic relationship among them. There have been only limited applications of such relation extraction in marketing, primarily due to the computational and linguistic complexities involved in accurately making such relational inferences from unstructured data (see, e.g., the diabetes drugs application in Netzer et al. [2012]). However, as the methodologies used to extract entity relations evolve, we expect this to be a promising direction for marketers to take.

The most commonly used approaches for relation extraction are handwritten relationship rules, supervised machine learning approaches, and a combination of these approaches. At the most basic level, the researcher could write a set of rules that describe the required relationship. An example of such a rule may be the co-occurrence of product (e.g., "Ford"), attribute (e.g., "oil consumption"), and problem (e.g., "excessive"). However, such approaches tend to require many handwritten rules and have low recall (they miss many relations) and thus are becoming less popular.

A more common approach is to train a supervised machine learning tool. This could be linguistic agnostic approaches (e.g., deep learning) or natural language processing (NLP) approaches that aim to understand the linguistic relationship in the sentence. Such an approach requires a relatively large training data set provided by human coders in which various relationships (e.g., sentiment) are observed. One readily available tool for NLP-based relationship extraction is the Stanford Sentence and Grammatical Dependency Parser (http://nlp.stanford.edu:8080/parser/). The tool identifies the grammatical role of different words in the sentence to identify their relationship. For example, to assign a sentiment to a particular attribute, the parser first identifies the presence of an emotion word and then, in cases where a subject is present, automatically assesses if there is a grammatical relationship (e.g., in the sentence "the hotel was very nice," the adjective "nice" relates to the subject "hotel"). As with many off-the-shelf tools, the validity of the tool for a specific relation extraction needs to be tested.

Finally, beyond the relations between words/entities within one document, text can also be investigated across documents (e.g., online reviews, academic articles). For example, a temporal sequence of documents or a portfolio of documents across a group or community of communicators can be examined for interdependencies (Ludwig et al. 2013, 2014).

Text Analysis Metrics

Early work in marketing has tended to summarize unstructured text with structured proxies for this data. For example, in online reviews, researchers have used volume (e.g., Godes and Mayzlin 2004; Moe and Trusov 2011); valence, often captured by numeric ratings that supplement the text (e.g., Godes and Silva 2012; Moe and Schweidel 2012; Ying, Feinberg and Wedel 2006); and variance, often captured using entropy-type measures (e.g., Godes and Mayzlin 2004). However, these quantifiable metrics often mask the richness of the text. Several common metrics are often used to quantify the text itself, as we explain next.

Count measures. Count measures have been used to measure the frequency of each entity's occurrence, entities' co-occurrence, or entities' relations. For example, when using dictionaries to evaluate sentiment or other categories, researchers often use the proportion of negative and/or positive words in the document, or the difference between the two (Berger and Milkman 2012; Borah and Tellis 2016; Pennebaker et al. 2015; Schweidel and Moe 2014; Tirunillai and Tellis 2014). The problem with simple counts is that longer documents are likely to include more occurrences of every entity. For that reason, researchers often focus on the proportions of words in the document that belong to a particular category (e.g., positive sentiment). The limitation of this simple measure is that some words are more likely to appear than others. For example, the word “laptop” is likely to appear in almost every review in corpora that is composed of laptop reviews.

Accuracy measures. When evaluating the accuracy of text measures relative to human-coded or externally validated documents, measures of recall and precision are often used. Recall is the proportion of entities in the original text that the text-mining algorithm was able to successfully identify (it is defined by the ratio of true positives to the sum of true positives and false negatives). Precision is the proportion of correctly identified entities from all entities identified (it is defined by the ratio of true positives to the sum of true positives and false positives). On their own, recall and precision measures are difficult to assess because an improvement in one often comes at the expense of the other. For example, if one defines that every entity in the corpora is a brand, recall for brands will be perfect (you will never miss a brand if it exists in the text), but precision will be very low (there will be many false positive identifications of a brand entity).

To create the balance between recall and precision, one can use the F1 measure—a harmonic mean of the levels of recall and precision. If the researcher is more concerned with false positives than false negatives (e.g., it is more important to identify positives than negatives), recall and precision can be weighted differently. Alternatively, for unbalanced data with high proportions of true or false in the populations, a receiver operating characteristics curve can be used to reflect the relationship between true positives and false positives, and the area under the curve is often used as a measure of accuracy.

Similarity measures. In some cases, the researcher is interested in measuring the similarity between documents (e.g., Ludwig et al. 2013). How similar is the language used in two advertisements? How different is a song from its genre? In such cases, measures such as linguistic style matching, similarity in topic use (Berger and Packard 2018), cosine similarity, and the Jaccard index (e.g., Toubia and Netzer 2017) can be used to assess the similarity between the text of two documents.

Readability measures. In some cases, the researcher is interested in evaluating the readability of the text. Readability can reflect the sophistication of the writer and/or the ability of the reader to comprehend the text (e.g., Ghose and Ipeirotis 2011). Common readability measures include the Flesch–Kincaid reading ease and the simple measure of gobbledygook (SMOG) measures. These measures often use metrics such as average number of syllables and average number of words per sentence to evaluate the readability of the text. Readability measures often grade the text on a 1–12 scale reflecting the U.S. school grade-level needed to comprehend the text. Common text-mining packages have built-in readability tools.

The Validity of Text-Based Constructs

While the availability of text has opened up a range of research questions, for textual data to provide value, one must be able to establish its validity. Both internal validity (i.e., does text accurately measure the constructs and the relationship between them?) and external validity (i.e., do the test-based findings apply to phenomena outside the study?) can be established in various ways (Humphreys and Wang 2017). Table 5 describes how the text analysis can be evaluated to improve different types of validity (Cook and Campbell 1979).

Internal Validity

Internal validity is often a major threat in the context of text analysis because the mapping between words and the underlying dimension the research aims to measure (e.g., psychological state and traits) is rarely straightforward and can vary across contexts and textual outlets (e.g., formal news vs. social media). In addition, given the relatively young field of automated text analysis, validation of many of the methods and constructs is still ongoing.

Accordingly, it is important to confirm the internal validity of the approach used. A range of methods can be adopted to ensure construct, concurrent, convergent, discriminant, and causal validity. In general, the approach for ensuring internal validity is to ensure that the text studied accurately reflects the theoretical concept or topic being studied, does so in a way that is congruent with prior literature, is discriminant from other related constructs, and provides ample and careful evidence for the claims of the research.

Construct validity. Construct validity (i.e., does the text represent the theoretical concept?) is perhaps the most important to address when studying text. Threats to construct validity occur when the text provides improper or misleading evidence of the construct. For instance, researchers often rely on existing standardized dictionaries to extract constructs to ensure that their work is comparable with other work. However, these dictionaries may not always fit the particular context. For example, extracting sentiment from financial reports using sentiment tools developed for day-to-day language may not be appropriate. Particularly when attempting to extract complex constructs (e.g., psychological states and traits, relationships between consumers and products, and even sentiment), researchers should attempt to validate the constructs on the specific application to ensure that what is being extracted from the text is indeed what they intended to extract. Construct validity can also be challenged when homonyms or other words do not accurately reflect what researchers think they do.

Strategies for addressing threats to construct validity require that researchers examine how the instances counted in the data connect to the theoretical concept(s) (Humphreys and Wang 2017). Dictionaries can also be validated using a saturation approach, pulling a subsample of coded entries and verifying with a hit rate of approximately 80% (Weber 2005). Another method is to use input from human coders, as is done to support machine learning applications (as previously discussed). For example, one can use Amazon Mechanical Turk workers to label phrases on a scale from “very negative” to “very positive” for sentiment analysis and then use these words to create a

Table 5. Text Analysis Validation Techniques.

Type of Validity	Validation Technique	Description of Method for Validation	References
Internal Validity			
Construct validity	Dictionary validation	After draft dictionary is created, pull 10% of the sample and calculate the hit rate. Measures such as hit rates, precision, and recall can be used to measure accuracy.	Weber (2005)
		Have survey participants rate words included in the dictionary. Based on this data, the dictionary can also be weighted to reflect the survey data.	Brysbaert, Warriner, and Kuperman (2014)[a]
		Have three coders evaluate the dictionary categories. If two of the three coders agree that the word is part of the category, include; if not, exclude. Calculate overall agreement.	Humphreys (2010); Pennebaker, Francis, and Booth (2001)[a]
	Saturation	Pull 10% of instances coded from the data and calculate the hit rate. Adjust word list until saturation reaches 80% hit rate.	Weber (2005)
Concurrent validity	Multiple dictionaries	Calculate and compare multiple textual measures of the same construct (e.g., multiple sentiment measures)	Hartmann et al. (2018)
	Comparison of topics	Compare with other topic models of similar data sets in other research (e.g., hotel reviews)	Mankad et al. (2016)[a]
Convergent validity	Triangulation	Look within text data for converging patterns (e.g., positive/e emotion correlates with known-positive attributes); apply Principle Components Analysis to show convergent groupings of words	Humphreys (2010); Kern et al. (2016)
	Multiple operationalizations	Operationalize constructs with textual and nontextual data (e.g., sentiment, star rating)	Ghose et al. (2012)[a]; Mudambi, Schuff, and Zhang (2014)[a]
Causal validity	Control variables	Include variables in the model that address rival hypotheses to control for these effects	Ludwig et al. (2013)
	Laboratory study	Replicate focal relationship between the independent variable and dependent variable in a laboratory setting	Spiller and Belogolova (2016)[a]; Van Laer et al. (2018)
External Validity			
Generalizability	Replication with different data sets	Compare the results from the text analysis with the results obtained other (possibly non-text-related) data sets	Netzer et al. (2012)
	Predict key performance measure	Include results from text analysis in regression or other model to predict a key outcome (e.g., sales, engagement)	Fossen and Schweidel (2019)
Predictive validity	Holdout sample	Train model on approximately 80%–90% of the data and validate the model with the remaining data. Validation can be done using k-fold validation, which trains the mode on k-1 subsets of the data and predicts for the remaining subset of testing.	Jurafsky et al. (2014)
Robustness	Different statistical measures, unitizations	Use different, but comparable, statistical measures or algorithms (e.g., lift, cosine similarity, Jaccard similarity), aggregate at different levels (e.g., day, month)	Netzer et al. (2012)

[a]Reference appears in the Web Appendix.

weighted dictionary. In many cases, multiple methods for dictionary validation are advisable to ensure that one is achieving both theoretical and empirical fit. For topic modeling, researchers infer topics from a list of cooccurring words. However, these are theoretical inferences made by researchers. As such, construct validity is equally important and can be ascertained using some of the same methods of validation, through saturation and calculating a hit rate through manual analysis of a subset of the data. When using a classification approach, confusion matrices can be produced to provide details on accuracy, false positives, and false negatives (Das and Chen 2007).

Concurrent validity. Concurrent validity concerns the way that the researcher's operationalization of the construct relates to prior operationalizations. Threats to concurrent validity often come when researchers create text-based measures inductively from the text. For instance, if one develops a topic model from the text, it will be based on the data set and may not therefore produce topics that are comparable with previous research. To address these threats, one should compare the operationalization with other research and other data sources. For example, Schweidel and Moe (2014) propose a measure of brand sentiment based on social media text data and validate it by comparing it with brand measures obtained through a traditional marketing research survey. Similarly, Netzer et al. (2012) compare the market structure maps derived from textual information with those derived from product switching and surveys, and Tirunillai and Tellis (2014) compare the topics they identify with those found in *Consumer Reports*. When studying linguistic style (Pennebaker and King 1999), for example, it is beneficial to use robust measures from prior literature where factor analysis and other methods have already been employed to create the construct.

Convergent validity. Convergent validity ensures that multiple measurements of the construct (i.e., words) all converge to the same concept. Convergent validity can be threatened when the measures of the construct do not align or have different effects. Convergent validity can be enhanced by using several substantively different measures (e.g., dictionaries) of the same construct to look for converging patterns. For example, when studying posts about the stock market, Das and Chen (2007) compare five classifiers for measuring sentiment, comparing them in a confusion matrix to examine false positives. Convergent evidence can also come from creating a correlation or similarity matrix of words or concepts and checking for patterns that have face validity. For instance, Humphreys (2010) looks for patterns between the concept of crime and negative sentiment to provide convergent evidence that crime is negatively valenced in the data.

Discriminant validity. Discriminant validity, the degree to which the construct measures are sufficiently different from measures of other constructs, can be threatened when the measurement of the construct is very similar to that of another construct. For instance, measurements of sentiment and emotion in many cases may not seem different because they are measured using similar word lists or, when using classification, return the same group of words as predictors. Strategies for ensuring discriminant validity entail looking for discriminant rather than convergent patterns and boundary conditions (i.e., when and how is sentiment different from emotion?). Furthermore, theoretical refinements can be helpful in drawing finer distinctions. For example, anxiety, anger, and sadness are different kinds of emotion (and can be measured via psychometrically different scales), whereas sentiment is usually measured as positive, negative, or neutral (Pennebaker et al. 2015).

Causal validity. Causal validity is the degree to which the construct, as operationalized in the data set, is actually the cause of another construct or outcome, and it is best ascertained through random assignment in controlled lab conditions. Any number of external factors can threaten causal validity. However, steps can be taken to enhance causal validity in naturally occurring textual data. In particular, rival hypotheses and other explanatory factors for the proposed causal relationship can be statistically controlled for in the model. For example, Ludwig et al. (2013) include price discount in the model when studying the relationship between product reviews and conversion rate to control for this factor.

External Validity

To achieve external validity, researchers should attempt to ensure that the effects found in the text apply outside of the research framework. Because text analysis often uses naturally occurring data that is often of large magnitude, it tends have a relatively high degree of external validity relative to, for example, lab experiments. However, establishing external validity is still necessary due to threats to validity from sampling bias, overfitting, and single-method bias. For example, online reviews may be biased due to self-selection among those who elected to review a product (Schoenmüller, Netzer, and Stahl 2019).

Predictive validity. Predictive validity is threatened when the construct, though perhaps properly measured, does not have the expected effects on a meaningful second variable. For example, if consumer sentiment falls but customer satisfaction remains high, predictive validity could be called into question. To ensure predictive validity, text-based constructs can be linked to key performance measures such as sales (e.g., Fossen and Schweidel 2019) or consumer engagement (Ashley and Tuten 2015). If a particular construct has been theoretically linked to a performance metric, then any text-based measure of that construct should also be linked to that performance metric. Tirunillai and Tellis (2012) show that the volume of Twitter activity affects stock price, but they find mixed results for the predictive validity of sentiment, with negative sentiment being predictive but positive sentiment having no effect.

Generalizability can be threatened when researchers base results on a single data set because it is unknown whether the findings, model, or algorithm would apply in the same way to other texts or outside of textual measurements. Generalizability of the results can be established by viewing the results of text analysis along with other measures of attitude and behavioral outcomes. For example, Netzer et al. (2012) test their substantive conclusions and methodology on message boards of both automobile discussions and drug discussions from WebMD. Evaluating the external validity and generalizability of the findings is key, because the analysis of text drawn from a particular source may not reflect consumers more broadly (e.g., Schweidel and Moe 2014).

Robustness. Robustness can be limited when there is only one metric or method used in the model. Researchers can ensure robustness by using different measures for relationships (e.g., Pearson correlation, cosine similarity, lift) and probing results by relaxing different assumptions. The use of holdout samples and k-fold cross-validation methods can prevent researchers from overfitting their models and ensure that relationships found in the data set will hold with other data as well (Jurafsky et al. 2014; see also Humphreys and Wang 2017). Probing on different “cuts” of the data can also help. Berger and Packard (2018), for example, compare lyrics from different genres, and Ludwig et al. (2013) include reviews of both fiction and nonfiction books.

Finally, researchers should bear in mind the limitations of text itself. There are thoughts and feelings that consumers, managers, or other stakeholders may not express in text. The form of communication (e.g., tweets, annual reports) may also shape the message; some constructs may not be explicit enough to be measured with automated text analysis. Furthermore, while textual information can often involve large samples, these samples may not be representative. Twitter users, for example, tend to be younger and more educated (Smith and Anderson 2018). Those who contribute textual information, particularly in social media, may represent polarized points of view. When evaluating cultural products or social media, one should consider the system in which they are generated. Often viewpoints are themselves filtered through a cultural system (Hirsch 1986; McCracken 1988) or elevated by an algorithm, and the products make it through this process may share certain characteristics. For this reason, researchers and firms should use caution when making attributions on the basis of a cultural text. It is not necessarily a reflection of reality (Jameson 2005) but rather may represent ideals, extremes, or institutionalized perceptions, depending on the context.

Future Research Agenda

We hope this article encourages more researchers and practitioners to think about how they can incorporate textual data into their research. Communication and linguistics are at the core of studying text in marketing. Automated text analysis opens the black box of interactions, allowing researchers to directly access what is being said and how it is said in marketplace communication. The notion of text as indicative of meaning-making processes creates fascinating and truly novel research questions and challenges. There are many methods and approaches available, and there is no space to do all of them justice. While we have discussed several research streams, given the novelty of text analysis, there are still ample opportunities for future research, which we discuss next.

Using Text to Reach Across the Marketing Discipline

Returning to how text analysis can unite the tribes of marketing, it is worth highlighting a few areas that have mostly been examined by one research tradition in marketing where fruitful cross-pollination between tribes is possible through text analysis. Brand communities were first identified and studied by researchers coming from a sociology perspective (Muñiz and O’Guinn 2001). Later, qualitative and quantitative researchers further refined the concepts, identifying a distinct set of roles and status in the community (e.g., Mathwick, Wiertz, and De Ruyter 2007). Automated text analysis allows researchers to study how consumers in these communities interact at scale and in a more quantifiable manner—for instance, examining how people with different degrees of power use language and predict group outcomes based on quantifiably different dynamics (e.g., Manchanda, Packard, and Pattabhitamaiah 2015). Researchers can track influence, for example, by investigating which types of users initiate certain words or phrases and which others pick up on them. Research could examine whether people begin to enculturate to the language of the community over time and predict which individuals may be more likely to stay or leave on the basis of how well they adapt to the group’s language (Danescu-Niculescu-Mizil et al. 2013; Srivastava and Goldberg 2017). Quantitative or machine learning researchers might capture the most commonly discussed topics and how these dynamically change over the evolution of the community. Interpretive researchers might examine how these terms link conceptually, to find underlying community norms that lead members to stay. Marketing strategy researchers might then use or develop dictionaries to connect these communities to firm performance and to offer directions for firms regarding how to keep members participating across different brand communities (or contexts).

The progression can flow the other way as well. Outside of a few early investigations (e.g., Dichter 1966), word of mouth was originally studied by quantitative researchers interested in whether interpersonal communication actually drove individual and market behavior (e.g., Chevalier and Mayzlin 2006; Iyengar, Van den Bulte, and Valente 2011). More recently, however, behavioral researchers have begun to study the underlying drivers of word of mouth, looking at why people talk about and share some stories, news, and information rather than others (Berger and Milkman 2012; De Angelis et al. 2012; for a review, see Berger [2014]). Marketing strategy researchers might track the text of word-of-mouth interactions to predict the emergence of brand crises or social media firestorms (e.g., Zhong and Schweidel 2019) as well as when, if, and how to respond (Herhausen et al. 2019).

Consumer–firm interaction is also a rich area to examine. Behavioral researchers could use the data from call centers to better understand interpersonal communication between consumers and firms and record what drives customer satisfaction (e.g., Packard and Berger 2019; Packard, Moore, and McFerran 2018). The back-and-forth between customers and agents could be used to understand conversational dynamics. More quantitative researchers should use the textual features of call centers to predict outcomes such as churn and even go beyond text to examine vocal features such as tone, volume, and speed of speech. Marketing strategy researchers could use calls to understand how customer-centric a company is or assess the quality, style, and impact of its sales personnel.

Finally, it is worth noting that different tribes not only have different skill sets but also often study substantively different types of textual communication. Consumer-to-consumer communication is often studied by researchers in consumer behavior, whereas marketing strategy researchers more often tend to study firm-to-consumer and firm-to-firm communication. Collaboration among researchers from the different subfields may allow them to combine these different sources of textual data.

There is ample opportunity to apply theory developed in one domain to enhance another. Marketing strategy researchers, for example, often use transaction economics to study business-to-business relationships through agency theory, but these approaches may be equally beneficial when studying consumer-to-consumer communications.

Broadening the Scope of Text Research

As noted in Table 1, certain text flows have been studied more than others. A large portion of existing work has focused on consumers communicating to one another through social media and online reviews. The relative availability of such data has made it a rich area of study and an opportunity to apply text analysis to marketing problems.[3] Furthermore, for this area to grow, researchers need to branch out. This includes expanding (1) data sources, (2) actors examined, and (3) research topics.

Expand data sources used. Offline word of mouth, for example, can be examined to study what people talk about and conversational dynamics. Doctor–patient interactions can be studied to understand what drives medical adherence. Text items such as yearbook entries, notes passed between students, or the text of speed dating conversations can be used to examine relationship formation, maintenance, and dissolution. Using offline data requires carefully transcribing content, which increases the amount of effort required but opens up a range of interesting avenues of study. For example, we know very little about the differences between online recommendations and face-to-face recommendations, where the latter also include the interplay between verbal and nonverbal information. Moreover, in the new era of "perpetual contact" our understanding of cross-message and cross-channel implications is limited. Research by Batra and Keller (2016) and Villarroel Ordenes et al. (2018) suggests that appropriate sequencing of messages matters; it might similarly matter across channels and modality. Given the rise of technology-enabled realities (e.g., augmented reality, virtual reality, mixed reality), assistive robotics, and smart speakers, understanding the roles and potential differences between language and nonverbal cues could be achieved using these novel data sources.

Expand dyads between text producers and text receivers. There are numerous dyads relevant to marketing in which text plays a crucial role. We discuss just a few of the areas that deserve additional research.

Considering consumer–firm interactions, we expect to see more research leveraging the rich information exchanged between consumers and firms through call centers and chats (e.g., Packard and Berger 2019; Packard, Moore, and McFerran 2018). These interactions often reflect inbound communication between customers and the firm, which can have important implications for the relationship between parties. In addition, how might the language used on packaging or in brand mission statements reflect the nature of organizations and their relationship to their consumers? How might the language that is most impactful in sales interactions differ from the language that is most useful in customer service interactions? Research could also probe how the impact of such language varies across contexts. The characteristics of language used by consumer packaged goods brands and pharmaceuticals brands in direct-to-consumer advertising likely differ. Similarly, the way in which consumers process the language used in disclosures in advertisements for pharmaceuticals (e.g., Narayanan, Desiraju, and Chintagunta 2004) and political candidates (e.g., Wang, Lewis, and Schweidel 2018) may vary.

Turning to firm-to-firm interactions, most conceptual frameworks on business-to-business (B2B) exchange relations emphasize the critical role of communication (e.g., Palmatier, Dant, and Grewal 2007). Communicational aspects have been linked to important B2B relational measures such as commitment, trust, dependence, relationship satisfaction, and relationship quality. Yet research on actual, word-level B2B communication is very limited. For example, very little research has examined the types of information exchanged between salespeople and customers in offline settings. The ability to gather and transcribe data at scale points to important opportunities to do so. As for within-firm communication, researchers could study informal communications such as marketing-related emails, memos, and agendas generated by firms and consumed by their employees.

Similarly, while a great deal of work in accounting and finance has begun to use annual reports as a data source (for a review, see Loughran and McDonald [2016]), marketing researchers have paid less attention to this area to study communication with investors. Most research has used this data to predict outcomes such as stock performance and other measures of firm valuation. Given recent interest in linking marketing-related activities to firm valuation (e.g., McCarthy and Fader 2018), this may be an area to pursue further. All firm communication, including required documents such as annual reports or discretionary forms of communication such as advertising and sales interactions, can be used to measure variables such as market orientation, marketing capabilities, marketing leadership styles, and even a firm's brand personality.

There are also ample research opportunities in the interactions between consumers, firms, and society. Data about the broader cultural and normative environment of firms, such as news media and government reports, may be useful to shed light on the forces that shape markets. To understand how a company such as Uber navigates resistance to market change, for example, one might study transcripts of town hall meetings and other government documents in which citizen input is heard and answered. Exogenous shocks in the forms of social movements such as #metoo and #blacklivesmatter have affected marketing communication and brand image. One potential avenue for future research is to take a cultural

[3] While readily available data facilitates research, there are downsides to be recognized, including the representatives of such data and the terms of service that govern the use of this data.

branding approach (Holt 2016) to study how different publics define, shape, and advocate for certain meanings in the marketplace. Firms and their brands do not exist in a vacuum, independent of the society in which they operate. Yet limited research in marketing has considered how text can be used to derive firms' intentions and actions at the societal level. For example, scholars have shown how groups of consumers such as locavores (i.e., people who eat locally grown food; Thompson and Coskuner-Balli 2007), fashionistas (Scaraboto and Fischer 2012), and bloggers (McQuarrie, Miller, and Phillips 2012) shape markets. Through text analysis, the effect of the intentions of these social groups on the market can then be measured and better understood.

Another opportunity for future research is the use of textual data to study culture and cultural success. Topics such as cultural propagation, artistic change, and the diffusion of innovations have been examined across disciplines with the goal of understanding why certain products succeed while others fail (Bass 1969; Boyd and Richerson 1986; Cavalli-Sforza and Feldman 1981; Rogers 1995; Salganik, Dodds, and Watts 2006; Simonton 1980). While success may be random (Bielby and Bielby 1994; Hirsch 1972), another possibility is that cultural items succeed or fail on the basis of their fit with consumers (Berger and Heath 2005). By quantifying aspects of books, movies, or other cultural items quickly and at scale, researchers can measure whether concrete narratives are more engaging, whether more emotionally volatile movies are more successful, whether songs that use certain linguistic features are more likely to top the Billboard charts, and whether books that evoke particular emotions sell more copies. While not as widely available as social media data, more and more data on cultural items has recently become available. Data sets such as the Google Books corpus (Akpinar and Berger 2015), song lyric websites, or movie script databases provide a wealth of information. Such data could enable analyses of narrative structure to identify "basic plots" (e.g., Reagan et al. 2016; Van Laer et al. 2019).

Key Marketing Constructs (That Could Be) Measured with Text

Beginning with previously developed ways of representing marketing constructs can help some researchers address validity concerns. This section details a few of these constructs to aid researchers who are beginning to use text analysis in their work (see the Web Appendix). Using prior operationalization of a construct can ensure concurrent validity—helping build the literature in a particular domain—but researchers should take steps to ensure that the prior operationalization has construct validity with their data set.

At the individual level, sentiment and satisfaction are perhaps some of the most common measurements (e.g., Büschken and Allenby, 2016; Homburg, Ehm, and Artz 2015; Herhausen et al. 2019; Ma, Baohung, and Kekre 2015; Schweidel and Moe 2014) and have been validated in numerous contexts. Other aspects that may be extracted from text include the authenticity and emotionality of language, which have also been explored through robust surveys and scales or by combining multiple existing measurements (e.g., Mogilner, Kamvar, and Aaker 2011; Van Laer et al. 2019). There are also psychological constructs, such as personality type and construal level (Kern et al. 2016; Snefjella and Kuperman 2015), that are potentially useful for marketing researchers and could also be inferred from the language used by consumers.

Future work in marketing studying individuals might consider measurements of social identification and engagement. That is, researchers currently have an idea of positive or negative consumer sentiment, but they are only beginning to explore emphasis (e.g., Rocklage and Fazio 2015), trust, commitment, and other modal properties. To this end, harnessing linguistic theory of pragmatics and examining phatics over semantics could be useful (see, e.g., Villarroel et al. 2017). Once such work is developed, we recommend that researchers carefully validate approaches proposed to measure such constructs along the lines described previously.

At the firm level, constructs have been identified in firm-produced text such as annual reports and press releases. Market orientation, advertising goals, future orientation, deceitful intentions, firm focus, and innovation orientation have all been measured and validated using this material (see Web Appendix Table 1). Work in organizational studies has a history of using text analysis in this area and might provide some inspiration and validation in the study of the existence of managerial frames for sensemaking and the effect of activists on firm activities.

Future work in marketing at the firm level could further refine and diversify measurements of strategic orientation (e.g., innovation orientation, market-driving vs. market-driven orientations). Difficult-to-measure factors deep in the organizational culture, structure, or capabilities may be revealed in the words the firm, its employees, and external stakeholders use to describe it (see Molner, Prabhu, and Yadav [2019]). Likewise, the mindsets and management style of marketing leaders may be discerned from the text they use (see Yadav, Prabhu, and Chandy [2007]). Firm attributes that are important outcomes of firm action (e.g., brand value) could also be explored using text (e.g., Herhausen et al. 2019). In this case, there is an opportunity to use new kinds of data. For instance, internal, employee-based brand value could be measured with text on LinkedIn or Glassdoor. Finally, more subtle attributes of firm language, including conflict, ambiguity, or openness, might provide some insight into the effects of managerial language on firm success. For this, it may be useful to examine less formal textual data of interactions such as employee emails, salesperson calls, or customer service center calls.

Less work in marketing has measured constructs on the social or cultural level, but work in this vein tends to focus on how firms fit into the cultural fabric of existing meanings and norms. For instance, institutional logics and legitimacy have been measured by analyzing media text, as has the rise of brand publics that increase discussion of brands within a culture (Arvidsson and Caliandro 2016).

At the cultural level, marketing research is likely to maintain a focus on how firms fit into the cultural environment, but it may also look to how the cultural environment affects consumers. For instance, measurement of cultural uncertainty, risk, hostility, and change could benefit researchers interested in the effects of culture on both consumer and firm effects as well as the effects of culture and society on government and investor relationships. Measuring openness and diversity through text are also timely topics to explore and might inspire innovations in measurement, focusing on, for example, language diversity rather than the specific content of language. Important cultural discourses such as language around debt and credit could also be better understood through text analysis. Measurement of gender- and race- related language could be useful in exploring diversity and inclusion in the way firms and consumers react to text from a diverse set of writers.

Opportunities and Challenges Provided by Methodological Advances

Opportunities. As the development of text analysis tools advances, we expect to see new and improved use of these tools in marketing, which can enable scholars to answer questions we could not previously address or have addressed only in a limited manner. Here are a few specific method-driven directions that seem promising.

First, the vast majority of the approaches used for text analysis in marketing (and elsewhere) rely on bag-of-words approaches, and thus, the ability to capture true linguistic relationships among words beyond their cooccurrence was limited. However, in marketing we are often interested in capturing the relationship among entities. For example, what problems or benefits did the customer mention about a particular feature of a particular product? Such approaches require capturing a deeper textual relationship among entities than is commonly used in marketing. We expect to see future development in these areas as deep learning and NLP-based approaches enable researchers to better capture semantic relationships.

Second, in marketing we are often interested in the latent intention or latent states of writers when creating text, such as their emotions, personality, and motivations. Most of the research in this area has relied on a limited set of dictionaries (primarily the LIWC dictionary) developed and validated to capture such constructs. However, these dictionaries are often limited in capturing nuanced latent states or latent states that may manifest differently across contexts. Similar to advances made in areas such as image recognition, with the availability of a large number of human-coded training data (often in the millions) combined with deep learning tools, we hope to see similar approaches being taken in marketing to capture more complex behavioral states from text. This would require an effort to human-code a large and diverse set of textual corpora for a wide range of behavioral states. Transfer learning methods commonly used in deep learning tools such as conventional neural nets can then be used to apply the learning from the more general training data to any specific application.

Third, there is also the possibility of using text analysis to personalize customer–firm interactions. Using machine learning, text analysis can also help personalize the customer interaction by detecting consumer traits (e.g., personality) and states (e.g., urgency, irritation) and perhaps eventually predicting traits associated with value to the firm (e.g., customer lifetime value). After analysis, firms can then tailor customer communication to match linguistic style and perhaps funnel consumers to the appropriate firm representative. The stakes of making such predictions may be high, mistakes costly, and there are clearly contexts in which using artificial intelligence impedes constructing meaningful customer–firm relationships (e.g., health care; Longoni, Bonezzi, and Morewedge 2019).

Fourth, while our discussion has focused on textual content, text is just one example of unstructured data, with audio, video, and image being others. Social media posts often marry text with images or videos. Print advertising usually overlays text on a carefully constructed visual. Although television advertising may not include text on the screen, it may have an audio track that contains text that progresses simultaneously with the video.

Until recently, text data has received the most attention, mainly due to the presence of tools to extract meaningful features. That said, tools such as Praat (Boersma 2001) allow researchers to extract information from audio (e.g., Van Zant and Berger 2019). One of the advantages of audio data over text data is that it provides richness in the form of tone and voice markers that can add to the actual words expressed (e.g., Xiao, Kim, and Ding 2013). This enables researchers to study not just what was said, but how it was said, examining how pitch, tone, and other vocal or paralinguistic features shape behavior.

Similarly, recent research has developed approaches to analyze images (e.g., Liu, Xuan et al. 2018), either characterizing the content of the image or identifying features within an image. Research into the impact of the combination of text and images is sparse (e.g., Hartmann et al. 2019). For example, images can be described in terms of their colors. In the context of print advertising, textual content may be less persuasive when used in conjunction with images of a particular color palette, whereas other color palettes may enhance the persuasiveness of text. Used in conjunction with simple images, the importance of text may be quite pronounced. But, when text is paired with complex imagery, viewers may attend primarily to the image, diminishing the impact of the text. If this is the case, legal disclosures that are part of an advertisement's fine print may not attract the audience's attention.

Analogous questions arise as to the role that text plays when incorporated into videos. Research has proposed approaches to characterize video content (e.g., Liu et al. 2018). In addition to comprising the script of the video, text may also appear visually. In addition to the audio context in which text appears, its impact may depend on the visuals that appear simultaneously. It may also be the case that its position within a video, relative to the start of the video, may moderate its effectiveness. For example, emotional text content that is spoken later in a video may be less persuasive for several reasons (e.g., the audience may have ceased paying attention by the time the text is spoken). Alternatively, the visuals with which the audio is paired may be more compelling to viewers, or the previous content of the video may have depleted a viewer's attentional resources. As our discussion of both images and videos suggests, text is but one component of marketing communications. Future research must investigate its interplay with other characteristics, including not only the content in which it appears but also when it appears (e.g., Kanuri, Chen, and Sridhar 2018), and in what media.

Challenges. While there are a range of opportunities, textual data also brings with it various challenges. First is the interpretation challenge. In some ways, text analysis seems to provide more objective ways of measuring behavioral processes. Rather than asking people how much they focused on themselves versus others when sharing word of mouth, for example, one can count the number of first-person (e.g., "I") and second-person (e.g., "you"; Barasch and Berger 2014) pronouns, providing what seems more like ground truth. But while part of this process is certainly more objective (e.g., the number of different types of pronouns), the link between such measures and underlying processes (i.e., what it says about the word-of-mouth transmitter) still requires some degree of interpretation. Other latent modes of behavior are even more difficult to count. While some words (e.g., "love") are generally positive, for example, how positive they are may depend heavily on idiosyncratic individual differences as well as the context.

More generally, there is challenge and opportunity in understanding the context in which textual information appears. While early work in this space, particularly research using entity extraction, asked questions such as how much emotion is in a passage of text, more accurate answers to that question take must take context into account. A restaurant review may contain lots of negative words, for example, but does that mean the person hates the food, the service, or the restaurant more generally? Songs that contain more second person-pronouns (e.g., "you") may be more successful (Packard and Berger 2019), but to understand why, it helps to know whether the lyrics use "you" as the subject or object of the sentence. Context provides meaning, and the more one understands not just which words are being used but also how they are being used, the easier it will be to extract insight. Dictionary-based tools are particularly susceptible to variation in the context in which the text appears, as dictionaries are often created in a context-free environment to match multiple contexts. Whenever possible, it is advised to use a dictionary that was created for the specific context of study (e.g., the financial sentiment tool developed by Loughran and McDonald [2016]).

As mentioned previously, there are also numerous methodological challenges. Particularly when exploring the "why," hundreds of features can be extracted, making it important to think about multiple hypothesis testing (and use of Bonferroni and other corrections). Only the text used by the text creator is available, so in some sense there is self-selection. Both the individuals who decide to contribute and the topics people decide to raise in their writing may suffer from self-selection. Particularly when text is used to measure (complex) behavioral constructs, validity of the constructs needs to be considered. In addition, for most researchers, analyzing textual information requires retooling and learning a whole new set of skills.

Data privacy challenges represent a significant concern. Research often uses online product reviews and sales ranking data scraped from websites (e.g., Wang, Mai, and Chiang 2013) or consumers' social media activity scraped from the platform (e.g., Godes and Mayzlin 2004; Tirunillai and Tellis 2012). Although such approaches are common, legal questions have started to arise. LinkedIn was unsuccessful in its attempt to block a startup company from scraping data that was posted on users' public profiles (Rodriguez 2017). While scraping public data may be permissible under the law, it may conflict with the terms of service of those platforms that have data of interest to researchers. For example, Facebook deleted accounts of companies that violated its data-scraping policies (Nicas 2018).[4] Such decisions raise important questions about the extent to which digital platforms can control access to content that users have chosen to make publicly available.

As interest in extracting insights from digitized text and other forms of digitized content (e.g., images, videos) grows, researchers should ensure that they have secured the appropriate permissions to conduct their work. Failure to do so may result in it becoming more difficult to conduct such projects. One potential solution is the creation of an academic data set, such as that made available by Yelp (https://www.yelp.com/dataset), which may contain outdated or scrubbed data to ensure that it does not pose any risk to the company's operations or user privacy.

The collection and analysis of digitized text, as well as other user-created content, also raises questions around users' expectations for privacy. In the wake of the European Union's General Data Protection Regulation and revelations about Cambridge Analytica's ability to collect user data from Facebook, researchers must be mindful of the potential abuses of their work. We should also consider the extent to which we are overstepping the intended use of user-generated content. For example, while a user may understand that actions taken on Facebook may result in their being targeted with specific advertisements for brands with which they have interacted, they may not anticipate the totality of their Facebook and Instagram activity being used to construct psychographic profiles that may be used by other brands. Understanding consumers' privacy preferences with regard to their online behaviors and the text they make available could provide important guidance for practitioners and researchers alike. Another rich area for future research is the advancement of the precision with which marketing can be implemented while minimizing intrusions of privacy (e.g., Provost et al. 2009).

[4] Facebook's terms of service with regard to automated data collection can be found at https://www.facebook.com/apps/site_scraping_tos_terms.php.

Concluding Thoughts

Communication is an important facet of marketing that encompasses communication between organizations and their partners, between businesses and their consumers, and among consumers. Textual data holds details of these communications, and through automated textual analysis, researchers are poised to convert this raw material into valuable insights. Many of the recent advances in the use of textual data were developed in fields outside of marketing. As we look toward the future and the role of marketers, these recent advancements should serve as exemplars. Marketers are well positioned at the interface between consumers, firms, and organizations to leverage and advance tools to extract textual information to address some of the key issues faced by business and society today, such as the proliferation of misinformation, the pervasiveness of technology in our lives, and the role of marketing in society. Marketing offers an invaluable perspective that is vital to this conversation, but it will only be by taking a broader perspective, breaking theoretical and methodological silos, and engaging with other disciplines that our research can reach its largest possible audience to affect the public discourse. We hope this framework encourages a reflection on the boundaries that have come to define marketing and opens avenues for future groundbreaking insights.

Editors

Christine Moorman and Harald van Heerde

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

Akpinar, Ezgi, and Jonah Berger (2015), "Drivers of Cultural Success: The Case of Sensory Metaphors," *Journal of Personality and Social Psychology*, 109 (1), 20–34.

Alessa, Ali, and Miad Faezipour (2018), "A Review of Influenza Detection and Prediction Through Social Networking Sites," *Theoretical Biology and Medical Modelling*, 15 (1), 1–27.

Anderson, Eric T., and Duncan I. Simester (2014), "Reviews Without a Purchase: Low Ratings, Loyal Customers, and Deception," *Journal of Marketing Research*, 51 (3), 249–69.

Arsel, Zeynep, and Jonathan Bean (2013), "Taste Regimes and Market-Mediated Practice," *Journal of Consumer Research*, 39 (5), 899–917.

Arvidsson, Adam, and Alessandro Caliandro (2016), "Brand Public," *Journal of Consumer Research*, 42 (5), 727–48.

Ashley, Christy, and Tracy Tuten (2015), "Creative Strategies in Social Media Marketing: An Exploratory Study of Branded Social Content and Consumer Engagement," *Psychology & Marketing*, 32 (1), 15–27.

Barasch, Alixandra, and Jonah Berger (2014), "Broadcasting and Narrowcasting: How Audience Size Affects What People Share," *Journal of Marketing Research*, 51 (3), 286–99.

Bass, Frank M. (1969), "A New Product Growth for Model Consumer Durables," *Management Science*, 15 (5), 215–27.

Batra, Rajeev, and Kevin L. Keller (2016), "Integrating Marketing Communications: New Findings, New Lessons, and New Ideas," *Journal of Marketing*, 80 (6), 122–45.

Berger, Jonah (2014), "Word of Mouth and Interpersonal Communication: A Review and Directions for Future Research," *Journal of Consumer Psychology*, 24 (4), 586–607.

Berger, Jonah, and Chip Heath (2005), "Idea Habitats: How the Prevalence of Environmental Cues Influences the Success of Ideas," *Cognitive Science*, 29 (2), 195–221.

Berger, Jonah, Yoon Duk Kim, and Robert Meyer (2019), "Emotional Volatility and Cultural Success," working paper.

Berger, Jonah, and Katherine L. Milkman (2012), "What Makes Online Content Viral?" *Journal of Marketing Research*, 49 (2), 192–205.

Berger, Jonah, Wendy W. Moe, and David A. Schweidel (2019), "What Makes Stories More Engaging? Continued Reading in Online Content," working paper.

Berger, Jonah, and Grant Packard (2018), "Are Atypical Things More Popular?" *Psychological Science*, 29 (7), 1178–84.

Berman, Ron, Shiri Melumad, Colman Humphrey, and Robert Meyer (2019), "A Tale of Two Twitterspheres: Political Microblogging During and After the 2016 Primary and Presidential Debates," *Journal of Marketing Research*, 56 (6), doi:10.1177/00222437 19861923.

Bielby, William, and Denise Bielby (1994), "'All Hits Are Flukes': Institutionalized Decision Making and the Rhetoric of Network Prime-Time Program Development," *American Journal of Sociology*, 99 (5), 1287–1313.

Blei, David M., Andrew Y. Ng, and Michael I. Jordan (2003), "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, 3, 993–1022.

Boersma, Paul (2001), "Praat, a System for Doing Phonetics by Computer," *Glot International*, 5 (9/10), 341–45.

Boghrati, Reihane, and Jonah Berger (2019) "Quantifying 60 Years of Misogyny in Music," working paper.

Bollen, Johan, Huina Mao, and Xiaojun Zeng (2011), "Twitter Mood Predicts the Stock Market," *Journal of Computational Science*, 2 (1), 1–8.

Borah, Abhishek, and Gerard J. Tellis (2016), "Halo (Spillover) Effects in Social Media: Do Product Recalls of One Brand Hurt or Help Rival Brands?" *Journal of Marketing Research*, 53 (2), 143–60.

Boyd, Robert, and Peter Richerson (1986), *Culture and Evolutionary Process*. Chicago: University of Chicago Press.

Büschken, Joachim, and Greg M. Allenby (2016), "Sentence-Based Text Analysis for Customer Reviews," *Marketing Science*, 35 (6), 953–75.

Cavalli-Sforza, Luigi Luca, and Marcus W. Feldman (1981), *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton, NJ: Princeton University Press.

Chen, Zoey, and Nicholas H. Lurie (2013), "Temporal Contiguity and Negativity Bias in the Impact of Online Word of Mouth," *Journal of Marketing Research*, 50 (4), 463–76.

Chevalier, Judith A., and Dina Mayzlin (2006), "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of Marketing Research*, 43 (3), 345–54.

Cohn, Michael A., Matthias R. Mehl, and James W. Pennebaker (2004), "Linguistic Markers of Psychological Change Surrounding September 11, 2001," *Psychological Science*, 15 (10), 687–93.

Cook, Thomas D., and Donald Thomas Campbell (1979), *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin.

Danescu-Niculescu-Mizil, Christian, Robert West, Dan Jurafsky, Jure Leskovec, and Christopher Potts (2013), "No Country for Old Members: User Lifecycle and Linguistic Change in Online Communities," in *Proceedings of the 22nd International Conference on World Wide Web*. New York: Association for Computing Machinery, 307–18.

Das, Sanjiv, and Mike Y. Chen (2007), "Yahoo! for Amazon: Sentiment Extraction from Small Talk on the Web," *Management Science*, 53 (9), 1375–88.

De Angelis, Matteo, Andrea Bonezzi, Alessandro M. Peluso, Derek Rucker, and Michele Costabile (2012), "On Braggarts and Gossips: A Self-Enhancement Account of Word-of-Mouth Generation and Transmission," *Journal of Marketing Research*, 49 (4), 551–63.

Dichter, E. (1966), "How Word-of-Mouth Advertising Works," *Harvard Business Review*, 44 (6), 147–66.

Dodds, Peter Sheridan, Harris Kameron Decker, Isabel M. Kloumann, Catherine A. Bliss, and Christopher M. Danforth (2011), "Temporal Patterns of Happiness and Information in a Global Social Network: Hedonometrics and Twitter," *PLoS ONE*, 6 (12), e26752.

Dowling, Grahame R., and Boris Kabanoff (1996), "Computer-Aided Content Analysis: What Do 240 Advertising Slogans Have in Common?" *Marketing Letters*, 7 (1), 63–75.

Eliashberg, Jehoshua, Sam K. Hui, and Z. John Zhang (2007a), "From Story Line to Box Office: A New Approach for Green-Lighting Movie Scripts," *Management Science*, 53 (6), 881–93.

Eliashberg, Jehoshua, Sam K. Hui, and Z. John Zhang (2007b), "Assessing Box Office Performance Using Movie Scripts: A Kernel-Based Approach," *IEEE Transactions on Knowledge & Data Engineering*, 26 (11), 2639–48.

Feldman, Ronen, Oded Netzer, Aviv Peretz, and Binyamin Rosenfeld (2015), "Utilizing Text Mining on Online Medical Forums to Predict Label Change Due to Adverse Drug Reactions," in *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York: Association for Computing Machinery, 1779–88.

Fiss, Peer C., and Paul M. Hirsch (2005), "The Discourse of Globalization: Framing and Sensemaking of an Emerging Concept," *American Sociological Review*, 70 (1), 29–52.

Fossen, Beth L., and David A. Schweidel (2019), "Social TV, Advertising, and Sales: Are Social Shows Good for Advertisers?" *Marketing Science*, 38 (2), 274–95.

Gandomi, Amir, and Murtaza Haider (2015), "Beyond the Hype: Big Data Concepts, Methods, and Analytics," *International Journal of Information Management*, 35 (2), 137–44.

Garg, Nikhil, Londa Schiebinger, Dan Jurafsky, and James Zou (2018), "Word Embeddings Quantify 100 Years of Gender and Ethnic Stereotypes," *Proceedings of the National Academy of Sciences*, 115 (16), E3635–44.

Gebhardt, Gary F., Francis J. Farrelly, and Jodie Conduit (2019), "Market Intelligence Dissemination Practices," *Journal of Marketing*, 83 (3), 72–90.

Ghose, Anindya, and Panagiotis G. Ipeirotis (2011), "Estimating the Helpfulness and Economic Impact of Product Reviews: Mining Text and Reviewer Characteristics," *IEEE Transactions on Knowledge and Data Engineering*, 23 (10), 1498–1512.

Godes, David, and Dina Mayzlin (2004), "Using Online Conversations to Study Word-of-Mouth Communication," *Marketing Science*, 23 (4), 545–60.

Godes, David, and José C. Silva (2012), "Sequential and Temporal Dynamics of Online Opinion," *Marketing Science*, 31 (3), 448–73.

Goffman, Erving (1959), "The Moral Career of the Mental Patient," *Psychiatry*, 22 (2), 123–42.

Gopalan, Prem, Jake M. Hofman, and David M. Blei (2013), "Scalable Recommendation with Poisson Factorization," (accessed August 19, 2019), https://arxiv.org/abs/1311.1704.

Hancock, Jeffrey T., Lauren E. Curry, Saurabh Goorha, and Michael Woodworth (2007), "On Lying and Being Lied To: A Linguistic Analysis of Deception in Computer-Mediated Communication," *Discourse Processes*, 45 (1), 1–23.

Hartmann, Jochen, Mark Heitmann, Christina Schamp, and Oded Netzer (2019), "The Power of Brand Selfies in Consumer-Generated Brand Images," working paper.

Hartmann, Jochen, Juliana Huppertz, Christina Schamp, and Mark Heitmann (2018), "Comparing Automated Text Classification Methods," *International Journal of Research in Marketing*, 36 (1), 20–38.

Hennig-Thurau, Thorsten, Caroline Wiertz, and Fabian Feldhaus (2015), "Does Twitter Matter? The Impact of Microblogging Word of Mouth on Consumers' Adoption of New Movies," *Journal of the Academy of Marketing Science*, 43 (3), 375–94.

Herhausen, Dennis, Stephan Ludwig, Dhruv Grewal, Jochen Wulf, and Marcus Schögel (2019), "Detecting, Preventing, and Mitigating Online Firestorms in Brand Communities," *Journal of Marketing*, 83 (3), 1–21.

Hill, Vanessa, and Kathleen M. Carley (1999), "An Approach to Identifying Consensus in a Subfield: The Case of Organizational Culture," *Poetics*, 27 (1), 1–30.

Hirsch, Arnold R. (1986) "The Last 'Last Hurrah'," *Journal of Urban History*, 13 (1), 99–110.

Hirsch, Peer M. (1972) "Processing Fads and Fashions: An Organization-Set Analysis of Cultural Industry Systems," *American Journal of Sociology*, 77 (4), 639–59.

Holt, Douglas (2016), "Branding in the Age of Social Media," *Harvard Business Review*, 94 (3), 40–50.

Homburg, Christian, Laura Ehm, and Martin Artz (2015), "Measuring and Managing Consumer Sentiment in an Online Community Environment," *Journal of Marketing Research*, 52 (5), 629–41.

Huang, Karen, Michael Yeomans, Alison W. Brooks, Julia Minson, and Francesca Gino (2017), "It Doesn't Hurt to Ask: Question-

Asking Increases Liking," *Journal of Personality and Social Psychology*, 113 (3), 430–52.

Humphreys, Ashlee (2010), "Semiotic Structure and the Legitimation of Consumption Practices: The Case of Casino Gambling," *Journal of Consumer Research*, 37 (3), 490–510.

Humphreys, Ashlee, and Kathryn A. LaTour (2013), "Framing the Game: Assessing the Impact of Cultural Representations on Consumer Perceptions of Legitimacy," *Journal of Consumer Research*, 40 (4), 773–95.

Humphreys, Ashlee, and Rebecca Jen-Hui Wang (2017), "Automated Text Analysis for Consumer Research," *Journal of Consumer Research*, 44 (6), 1274–1306.

Hutto, Clayton J., and Eric Gilbert (2014), "VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text," in *Proceedings of the Eighth International Conference on Weblogs and Social Media*. Palo Alto, CA: Association for the Advancement of Artificial Intelligence.

Iyengar, Raghuram, Christopher Van den Bulte, and Thomas Valente (2011), "Opinion Leadership and Social Contagion in New Product Diffusion," *Marketing Science*, 30 (2), 195–212.

Jameson, Fredric (2005), *Archaeologies of the Future: The Desire Called Utopia and Other Science Fictions*. New York: Verso.

Jurafsky, Dan, Victor Chahuneau, Bryan R. Routledge, and Noah A. Smith (2014), "Narrative Framing of Consumer Sentiment in Online Restaurant Reviews," *First Monday*, 19 (4), https://firstmonday.org/ojs/index.php/fm/article/view/4944/3863.

Kanuri, Vamsi K., Yixing Chen, and Shrihari (Hari) Sridhar (2018), "Scheduling Content on Social Media: Theory, Evidence, and Application," *Journal of Marketing*, 82 (6), 89–108.

Kern, Margaret L., Gregory Park, Johannes C. Eichstaedt, H. Andrew Schwartz, Maarten Sap, and Laura K. Smith (2016), "Gaining Insights from Social Media Language: Methodologies and Challenges," *Psychological Methods*, 21 (4), 507–25.

Kübler, Raoul V., Anatoli Colicev, and Koen Pauwels (2017) "Social Media's Impact on Consumer Mindset: When to Use Which Sentiment Extraction Tool," *Marketing Science Institute Working Paper Series* 17-122-09.

Kulkarni, Dipti (2014), "Exploring Jakobson's 'Phatic Function' in Instant Messaging Interactions," *Discourse & Communication*, 8 (2), 117–36.

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton (2015), "Deep Learning," *Nature*, 521 (7553), 436–44.

Lee, Thomas Y., and Eric T. Bradlow (2011), "Automated Marketing Research Using Online Customer Reviews," *Journal of Marketing Research*, 48 (5), 881–94.

Li, Feng, and Timon C. Du (2011), "Who Is Talking? An Ontology-Based Opinion Leader Identification Framework for Word-of-Mouth Marketing in Online Social Blogs," *Decision Support Systems*, 51 (1), 190–97.

Liu, Jia, and Olivier Toubia (2018), "A Semantic Approach for Estimating Consumer Content Preferences from Online Search Queries," *Marketing Science*, 37 (6), 855–1052.

Liu, Liu, Daria Dzyabura, and Natalie Mizik (2018), "Visual Listening In: Extracting Brand Image Portrayed on Social Media," working paper.

Liu, Xuan, Savannah Wei Shi, Thales Teixeira, and Michel Wedel (2018), "Video Content Marketing: The Making of Clips," *Journal of Marketing*, 82 (4), 86–101.

Ljung, M. (2000), "Newspaper Genres and Newspaper English," in *English Media Texts Past and Present: Language and Textual Structure*, Friedrich Ungerer, ed. Philadelphia: John Benjamins Publishing, 131–50.

Longoni, Chiara, Andrea A. Bonezzi, and Carey K. Morewedge (2019), "Resistance to Medical Artificial Intelligence," *Journal of Consumer Research*, published online May 3, DOI: https://doi.org/10.1093/jcr/ucz013.

Loughran, Tim, and Bill McDonald (2016). "Textual Analysis in Accounting and Finance: A Survey," *Journal of Accounting Research*, 54 (4), 1187–1230.

Ludwig, Stephan, Ko De Ruyter, Mike Friedman, Elisabeth C. Bruggen, Martin Wetzels, and Gerard Pfann (2013), "More Than Words: The Influence of Affective Content and Linguistic Style Matches in Online Reviews on Conversion Rates," *Journal of Marketing*, 77 (1), 87–103.

Ludwig, Stephan, Ko De Ruyter, Dominik Mahr, Elisabeth C. Bruggen, Martin Wetzels, and Tom De Ruyck (2014), "Take Their Word for It: The Symbolic Role of Linguistic Style Matches in User Communities," *MIS Quarterly*, 38 (4), 1201–17.

Ludwig, Stephan, Tom Van Laer, Ko De Ruyter, and Mike Friedman (2016), "Untangling a Web of Lies: Exploring Automated Detection of Deception in Computer-Mediated Communication," *Journal of Management Information Systems*, 33 (2), 511–41.

Ma, Liye, Sun Baohung, and Sunder Kekre (2015), "The Squeaky Wheel Gets the Grease—An Empirical Analysis of Customer Voice and Firm Intervention on Twitter," *Marketing Science*, 34 (5), 627–45.

Manchanda, Puneet, Grant Packard, and A. Pattabhitamaiah (2015), "Social Dollars: The Economic Impact of Consumer Participation in a Firm-Sponsored Online Community," *Marketing Science*, 34 (3), 367–87.

Mathwick, Charla, Caroline Wiertz, and Ko De Ruyter (2007), "Social Capital Production in a Virtual P3 Community," *Journal of Consumer Research*, 34 (6), 832–49.

McCarthy, Daniel, and Peter Fader (2018), "Customer-Based Corporate Valuation for Publicly Traded Noncontractual Firms," *Journal of Marketing Research*, 55 (5), 617–35.

McCombs, Maxwell E., and Donald L. Shaw (1972), "The Agenda-Setting Function of Mass Media," *Public Opinion Quarterly*, 36 (2), 176–87.

McCracken, Grant (1988), *Qualitative Research Methods: The Long Interview*. Newbury Park, CA: SAGE Publications.

McQuarrie, Edward F., Jessica Miller, and Barbara J. Phillips (2012), "The Megaphone Effect: Taste and Audience in Fashion Blogging," *Journal of Consumer Research*, 40 (1), 136–58.

Melumad, Shiri J., J. Jeffrey Inman, and Michael Tuan Pham (2019), "Selectively Emotional: How Smartphone Use Changes User-Generated Content," *Journal of Marketing Research*, 56 (2), 259–75.

Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean (2013), "Efficient Estimation of Word Representations in Vector Space," (accessed August 19, 2019), https://arxiv.org/abs/1301.3781.

Moe, Wendy W., and David A. Schweidel (2012), “Online Product Opinions: Incidence, Evaluation, and Evolution,” *Marketing Science*, 31 (3), 372–86.

Moe, Wendy W., and Michael Trusov (2011), “The Value of Social Dynamics in Online Product Ratings Forums,” *Journal of Marketing Research*, 48 (3), 444–56.

Mogilner, Cassie, Sepandar D. Kamvar, and Jennifer Aaker (2011), “The Shifting Meaning of Happiness,” *Social Psychological and Personality Science*, 2 (4), 395–402.

Molner, Sven, Jaideep C. Prabhu, and Manjit S. Yadav (2019), “Lost in the Universe of Markets: Toward a Theory of Market Scoping for Early-Stage Technologies,” *Journal of Marketing*, 83 (2), 37–61.

Moody, Christopher E. (2016), “Mixing Dirichlet Topic Models and Word Embeddings to Make lda2vec,” (accessed August 19, 2019), https://arxiv.org/abs/1605.02019.

Moon, Sangkil, and Wagner A. Kamakura (2017), “A Picture Is Worth a Thousand Words: Translating Product Reviews into a Product Positioning Map,” *International Journal of Research in Marketing*, 34 (1), 265–85.

Moorman, Christine, Harald J. van Heerde, C. Page Moreau, and Robert W. Palmatier (2019a), “Challenging the Boundaries of Marketing,” *Journal of Marketing*, 83 (5), 1–4.

Moorman, Christine, Harald J. van Heerde, C. Page Moreau, and Robert W. Palmatier (2019b), “ *JM* as a Marketplace of Ideas,” *Journal of Marketing*, 83 (1), 1–7.

Muñiz, Albert, Jr., and Thomas O’Guinn (2001), “Brand Community,” *Journal of Consumer Research*, 27 (4), 412–32.

Narayanan, Sridhar, Ramarao Desiraju, and Pradeep K. Chintagunta (2004), “Return on Investment Implications for Pharmaceutical Promotional Expenditures: The Role of Marketing-Mix Interactions,” *Journal of Marketing*, 68 (4), 90–105.

Netzer, Oded, Ronen Feldman, Jacob Goldenberg, and Moshe Fresko (2012), “Mine Your Own Business: Market-Structure Surveillance Through Text Mining,” *Marketing Science*, 31 (3), 521–43.

Netzer, Oded, Alain Lemaire, and Michal Herzenstein (2019), “When Words Sweat: Identifying Signals for Loan Default in the Text of Loan Applications,” *Journal of Marketing Research*, published electronically August 15, 2019, doi:10.1177/0022243719852959.

Nicas, Jack (2018), “Facebook Says Russian Firms ‘Scraped’ Data, Some for Facial Recognition,” *The New York Times* (October 12), https://www.nytimes.com/2018/10/12/technology/facebook-russian-scraping-data.html.

Nisbett, Richard E., and Timothy D. Wilson (1977), “Telling More Than We Can Know: Verbal Reports on Mental Processes,” *Psychological Review*, 84 (3), 231–59.

Opoku, Robert, Russell Abratt, and Leyland Pitt (2006), “Communicating Brand Personality: Are the Websites Doing the Talking for the Top South African Business Schools?” *Journal of Brand Management*, 14 (1/2), 20–39.

Ott, Myle, Claire Cardie, and Jeff Hancock (2012), “Estimating the Prevalence of Deception in Online Review Communities,” in *Proceedings of the 21st International Conference on World Wide Web*. New York: Association for Computing Machinery, 201–10.

Packard, Grant, and Jonah Berger (2017), “How Language Shapes Word of Mouth’s Impact,” *Journal of Marketing Research*, 54 (4), 572–88.

Packard, Grant, and Jonah Berger (2019), “How Concrete Language Shapes Customer Satisfaction,” working paper.

Packard, Grant, Sarah G. Moore, and Brent McFerran (2018), “(I’m) Happy to Help (You): The Impact of Personal Pronoun Use in Customer–Firm Interactions,” *Journal of Marketing Research*, 55 (4), 541–55.

Palmatier, Robert W., Rajiv P. Dant, and Dhruv Grewal (2007), “A Comparative Longitudinal Analysis of Theoretical Perspectives of Interorganizational Relationship Performance,” *Journal of Marketing*, 71 (4), 172–94.

Peladeau, N. (2016), *WordStat: Content Analysis Module for SIMSTAT*. Montreal, Canada: Provalis Research.

Pennebaker, James W. (2011), “The Secret Life of Pronouns,” *New Scientist*, 211 (2828), 42–45.

Pennebaker, James W., Roger J. Booth, Ryan L. Boyd, and Martha E. Francis (2015), *Linguistic Inquiry and Word Count: LIWC2015*. Austin, TX: Pennebaker Conglomerates.

Pennebaker, James W., and Laura A. King (1999), “Linguistic Styles: Language Use as an Individual Difference,” *Journal of Personality and Social Psychology*, 77 (6), 1296–1312.

Smith, Aaron, and Monica Anderson (2018), “Social Media Use in 2018,” Pew Research Center (March 1), http://www.pewinternet.org/2018/03/01/social-media-use-in-2018/.

Pollach, Irene (2012), “Taming Textual Data: The Contribution of Corpus Linguistics to Computer-Aided Text Analysis,” *Organizational Research Methods*, 15 (2), 263–87.

Provost, Foster, Brian Dalessandro, Rod Hook, Xiaohan Zhang, and Alan Murray (2009), *“Audience Selection for On-Line Brand Advertising: Privacy-Friendly Social Network Targeting,”* in *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York: Association for Computing Machinery, 707–16.

Puranam, Dinesh, Vishal Narayan, and Vrinda Kadiyali (2017), “The Effect of Calorie Posting Regulation on Consumer Opinion: A Flexible Latent Dirichlet Allocation Model with Informative Priors,” *Marketing Science*, 36 (5), 726–46.

Ransbotham, Sam, Nicholas Lurie, and Hongju Liu (2019), “Creation and Consumption of Mobile Word of Mouth: How Are Mobile Reviews Different?” *Marketing Science*, published online January 28, DOI: https://doi.org/10.1287/mksc.2018.1115.

Reagan, Andrew J., Lewis Mitchell, Dilan Kiley, Christopher M. Danforth, and Petter Sheridan Dodds (2016), “The Emotional Arcs of Stories Are Dominated by Six Basic Shapes,” *EPJ Data Science*, 5 (1), 1–12.

Rocklage, Matthew D., and Russell H. Fazio (2015), “The Evaluative Lexicon: Adjective Use as a Means of Assessing and Distinguishing Attitude Valence, Extremity, and Emotionality,” *Journal of Experimental Social Psychology*, 56 (January), 214–27.

Rocklage, Matthew D., Derek D. Rucker, and Loran F. Nordgren (2018), “The Evaluative Lexicon 2.0: The Measurement of Emotionality, Extremity, and Valence in Language,” *Behavior Research Methods*, 50 (4), 1327–44.

Rodriguez, Salvador (2017), “U.S. Judge Says LinkedIn Cannot Block Startup from Public Profile Data,” *Reuters* (August 14), https://www.reuters.com/article/us-microsoft-linkedin-ruling/u-s-judge-says-linkedin-cannot-block-startup-from-public-profile-data-idUSKCN1AU2BV.

Rogers, Everett M. (1995). *Diffusion of Innovations*, 4th ed. New York: The Free Press.

Rosa, Jose 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.

Rude, Stephanie, Eva-Maria Gortner, and James Pennebaker (2004), "Language Use of Depressed and Depression-Vulnerable College Students," *Cognition & Emotion*, 18 (8), 1121–33.

Salganik, Matthew J., Peter Dodds, and Duncan Watts (2006), "Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market," *Science*, 311 (5762), 854–56.

Scaraboto, Daiane, and Eileen Fischer (2012), "Frustrated Fashionistas: An Institutional Theory Perspective on Consumer Quests for Greater Choice in Mainstream Markets," *Journal of Consumer Research*, 39 (6), 1234–57.

Schoenmüller, Verena, Oded Netzer, and Florian Stahl (2019), "The Extreme Distribution of Online Reviews: Prevalence, Drivers and Implications," Columbia Business School Research Paper.

Schweidel, David A., and Wendy W. Moe (2014), "Listening in on Social Media: A Joint Model of Sentiment and Venue Format Choice," *Journal of Marketing Research*, 51 (4), 387–402.

Simonton, Dean Keith (1980), "Thematic Fame, Melodic Originality, and Musical Zeitgeist: A Biographical and Transhistorical Content Analysis," *Journal of Personality and Social Psychology*, 38 (6), 972–83.

Snefjella, Bryor, and Victor Kuperman (2015), "Concreteness and Psychological Distance in Natural Language Use," *Psychological Science*, 26 (9), 1449–60.

Srivastava, Sameer B., and Amir Goldberg (2017), "Language as a Window into Culture," *California Management Review*, 60 (1), 56–69.

Stewart, David W., and David H. Furse (1986), *TV Advertising: A Study of 1000 Commercials*. Waltham, MA: Lexington Books.

Tausczik, Yla R., and James.W. Pennebaker (2010), "The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods," *Journal of Language and Social Psychology*, 29 (1), 24–54.

Tellis, Gerard J., Deborah J. MacInnis, Seshadri Tirunillai, and Yanwei Zhang (2019), "What Drives Virality (Sharing) of Online Digital Content? The Critical Role of Information, Emotion, and Brand Prominence," *Journal of Marketing*, 83 (4), 1–20.

Thompson, Craig J., and Gokcen Coskuner-Balli (2007), "Countervailing Market Responses to Corporate Co-Optation and the Ideological Recruitment of Consumption Communities," *Journal of Consumer Research*, 34 (2), 135–52.

Timoshenko, Artem, and John R. Hauser (2019), "Identifying Customer Needs from User-Generated Content," *Marketing Science*, 38 (1), 1–20.

Tirunillai, Seshadri, and Gerard J. Tellis (2012), "Does Chatter Really Matter? Dynamics of User-Generated Content and Stock Performance," *Marketing Science*, 31 (2), 198–215.

Tirunillai, Seshadri, and Gerard J. Tellis (2014), "Mining Marketing Meaning from Online Chatter: Strategic Brand Analysis of Big Data Using Latent Dirichlet Allocation," *Journal of Marketing Research*, 51 (4), 463–79.

Toubia, Olivier, Garud Iyengar, Renée Bunnell, and Alain Lemaire (2019), "Extracting Features of Entertainment Products: A Guided LDA Approach Informed by the Psychology of Media Consumption," *Journal of Marketing Research*, 56(1), 18–36.

Toubia, Olivier, and Oded Netzer (2017), "Idea Generation, Creativity, and Prototypicality," *Marketing Science*, 36 (1), 1–20.

Tsai, Jeanne L. (2007), "Ideal Affect: Cultural Causes and Behavioral Consequences," *Perspectives on Psychological Science*, 2 (3), 242–59.

Van Laer, Tom, Jennifer Edson Escalas, Stephan Ludwig, and Ellis A. Van den Hende (2018), "What Happens in Vegas Stays on TripAdvisor? Computerized Analysis of Narrativity in Online Consumer Reviews," *Journal of Consumer Research*, 46 (2), 267–85.

Van Zant, Alex B., and Jonah Berger (2019), "How the Voice Persuades," working paper, Rutgers University.

Villarroel Ordenes, Francisco, Dhruv Grewal, Stephan Ludwig, Ko De Ruyter, Dominik Mahr, and Martin Wetzels (2018), "Cutting Through Content Clutter: How Speech and Image Acts Drive Consumer Sharing of Social Media Brand Messages," *Journal of Consumer Research*, 45 (5), 988–1012.

Villarroel Ordenes, Francisco, Stephan Ludwig, Ko De Ruyter, Dhruv Grewal, and Martin Wetzels (2017), "Unveiling What Is Written in the Stars: Analyzing Explicit, Implicit, and Discourse Patterns of Sentiment in Social Media," *Journal of Consumer Research*, 43 (6), 875–94.

Vosoughi, Soroush, Deb Roy, and Sinan Aral (2018), "The Spread of True and False News Online," *Science*, 359 (6380), 1146–51.

Wang, Xin, Feng Mai, and Roger H.L. Chiang (2013), "Database Submission—Market Dynamics and User-Generated Content About Tablet Computers," *Marketing Science*, 33 (3), 449–58.

Wang, Yanwen, Michael Lewis, and David A. Schweidel (2018), "A Border Strategy Analysis of Ad Source and Message Tone in Senatorial Campaigns," *Marketing Science*, 37 (3), 333–55.

Weber, Klaus (2005), "A Toolkit for Analyzing Corporate Cultural Toolkits," *Poetics*, 33 (3/4), 227–52.

Wies, Simone, Arvid Oskar Ivar Hoffmann, Jaako Aspara, and Joost M.E. Pennings (2019), "Can Advertising Investments Counter the Negative Impact of Shareholder Complaints on Firm Value?" *Journal of Marketing*, 83 (4), 58–80.

Xiao, Li, Hye-Jin Kim, and Min Ding (2013), "An Introduction to Audio and Visual Research and Applications in Marketing," in *Review of Marketing Research*, Vol. 10, Naresh Malhotra, ed. Bingley, UK: Emerald Publishing, 213–53.

Xiong, Ying, Moonhee Cho, and Brandon Boatwright (2019), "Hashtag Activism and Message Frames Among Social Movement Organizations: Semantic Network Analysis and Thematic Analysis of Twitter During the #MeToo Movement," *Public Relations Review*, 45 (1), 10–23.

Yadav, Manjit S., Jaideep C. Prabhu, and Rajesh K. Chandy (2007), "Managing the Future: CEO Attention and Innovation Outcomes," *Journal of Marketing*, 71 (4), 84–101.

Ying, Yuanping, Fred Feinberg, and Michel Wedel (2006), "Leveraging Missing Ratings to Improve Online Recommendation Systems," *Journal of Marketing Research*, 43 (3), 355–65.

Zhong, Ning, and David A. Schweidel (2019), "Capturing Changes in Social Media Content: A Multiple Latent Changepoint Topic Model," working paper, Emory University.