Received: 25 July 2018 | Revised: 21 May 2019 | Accepted: 25 May 2019 | Published on: 6 August 2019

DOI: 10.1002/smj.3067

RESEARCH ARTICLE

STRATEGIC MANAGEMENT JOURNAL WILEY

Machine learning approaches to facial and text analysis: Discovering CEO oral communication styles

Prithwiraj Choudhury[1] | **Dan Wang**[2] | **Natalie A. Carlson**[2] | **Tarun Khanna**[1]

[1]Harvard Business School, Boston, Massachusetts

[2]Columbia Business School, New York, New York

Correspondence
Prithwiraj Choudhury, Harvard Business School, Boston, MA.
Email: pchoudhury@hbs.edu

Abstract

***Research Summary*:** We demonstrate how a novel synthesis of three methods—(a) unsupervised topic modeling of text data to generate new measures of textual variance, (b) sentiment analysis of text data, and (c) supervised ML coding of facial images with a cutting-edge convolutional neural network algorithm—can shed light on questions related to CEO oral communication. With videos and corresponding transcripts of interviews with emerging market CEOs, we use this synthesis of methods to discover five distinct *communication styles* that incorporate both verbal and nonverbal aspects of communication. Our data comprises interviews that represent *unedited* expressions and content, making them especially suitable as data sources for the measurement of an individual's communication style. We then perform a proof-of-concept analysis, correlating CEO communication styles to M&A outcomes, highlighting the value of combining text and videographic data to define styles. We also discuss the benefits of using our methods versus current research methods.

***Managerial Summary*:** CEOs spend most of their time communicating to investors, customers, and partners with the aim of influencing these various stakeholders. To what extent though does their effectiveness as leaders depend on a mixture of what they say and how they say it? We use cutting-edge machine learning approaches to measure a CEO's communication style, which can give clues about the major strategic decisions a CEO's firm must make. With a collection of video interviews with 61 organizational leaders from emerging markets, we use textual analysis and facial image expression recognition to code whether CEOs are "excitable," "stern," "dramatic," "rambling," and "melancholy" in their communication styles. As a proof-of-concept, we also show that CEOs who were more dramatic in expressing themselves were also less likely to oversee major acquisitions. Therefore, not only can CEO communication styles help predict a firm's ability to grow, adapt to change, and reallocate existing assets, styles can also be coded more intuitively by using our new method, representing a vast improvement over previous methods in both accessibility and interpretability.

KEYWORDS

CEO oral communication, image analysis, machine learning, managerial cognitive capability, topic modeling

1 | INTRODUCTION

With the advent of empirical techniques based on machine learning (ML), research in social sciences is arguably at an inflection point (Athey, 2019). Recent papers in economics—such as Mullainathan and Spiess (2017) and Kleinberg, Lakkaraju, Leskovec, Ludwig, and Mullainathan (2017)—have demonstrated the usefulness of empirical predictive techniques that build on ML concepts. Machine learning techniques have been shown to be particularly helpful in analyzing new sources of "big data" that previously have been underutilized for research, such as large textual archives (Antweiler & Frank, 2004) and repositories of images (Glaeser, Kominers, Luca, & Naik, 2018). More broadly, research across several fields within management has started to embrace big data and text/image mining tools (e.g., Arts, Cassiman, & Gomez, 2018; Kaplan & Vakili, 2015; Menon, Tabakovic, & Lee, 2018; Riedl et al., 2016). In this paper, we detail a novel synthesis of ML methods for coding textual data and facial expressions to shine light on CEO oral communication. In doing so, we attempt to advance the study of CEO oral communication by: (a) synthesizing multiple methods and data related to both verbal and nonverbal aspects of communication to generate measures for a CEO's *communication style*, and (b) demonstrating the benefit of using state-of-the-art methods (e.g., a convolutional neural network method to code facial images) vis-à-vis methods currently being used in the literature (e.g., videometrics driven by human coding).

We study CEO oral communication in response to the call made by Helfat and Peteraf (2015) to study verbal language and nonverbal communication, which they highlight as important components of managerial cognitive capabilities. Communicating well is one of the most important skills in the CEO toolkit. As Bandiera, Guiso, Prat, and Sadun (2018) argue, CEOs need to create organizational alignment, and this requires significant investment in communication across a broad variety of constituencies, including the persuasion of internal and external stakeholders to embrace cognitively distant opportunities (Gavetti, 2012). In prior research on CEO communication, the focus has been on content analysis of text from written communication by the CEO, using data such as CEO letters to stakeholders (Barr, 1998; Gamache & McNamara, 2019; Kaplan, 2008; Salancik & Meindl, 1984; Watzlawick, Bavelas, & Jackson, 1967); there is also a recent literature that analyzes transcripts of earnings conference call presentations (Pan, McNamara, Lee, Haleblian, & Devers, 2018). To code text-based communication, the current approach in CEO communication research is to use dictionary-based methods, such as the linguistic inquiry and word count (LIWC) software (Gamache & McNamara, 2019; Pan et al., 2018).

Helfat and Peteraf (2015) make a persuasive argument for why strategy scholars should study “oral language” (such as CEO oral communication) in addition to studying “language production” (such as CEO written communication). Both verbal and nonverbal aspects of CEO oral communication are related to “managerial cognitive capabilities” and rely less on controlled mental processing, compared to written communication. There is also nascent empirical scholarship in the strategy and accounting literatures analyzing CEO oral communication, that is, the analysis of what CEOs say (i.e., verbal oral communication), and their facial gestures (i.e., nonverbal oral communication) by studying videographic data (Blankespoor, Hendricks, & Miller, 2017; Hill, Petrenko, Ridge, & Aime, 2019; Petrenko, Aime, Ridge, & Hill, 2016). Here, the methodological tool of choice has been the “videometric” method, wherein third-party human raters are trained to code CEO expressions using psychometrically validated instruments. We instead employ a state-of-the-art convolutional neural network method to code facial expressions and use two methods to code CEO communication text: topic modeling based on unsupervised ML (the Latent Dirichlet Allocation or LDA model) and dictionary-based sentiment analysis.

Our first text-based method estimates unsupervised topic models through Latent Dirichlet Allocation (Blei, Ng, & Jordan, 2003). Topic modeling offers a systematic way of measuring the distribution of topics that describe the content of a set of documents in the form of sets of keywords (Kaplan & Vakili, 2015). Our second method of textual analysis relies on a more standard dictionary-based approach to conduct sentiment analysis. The sentiment measures in this paper are calculated using the *Syuzhet* R package (Jockers, 2017), which employs crowdsourced lexicons developed by Saif Mohammad at the National Resource Council Canada (NRC). Topic modeling and sentiment analysis enable us to analyze the content and valence of CEO oral communication. From these text-based methods, we calculate two novel measures: one indicating the *variance of sentiment* over the scope of the interview and another, *topic entropy*, indicating the diversity of semantic topics covered in the interview.

The third method employed in the paper uses supervised ML to code expressions of facial images. The underlying algorithm (to be explained in detail later) uses convolutional neural networks (Yu & Zhang, 2015) to code facial emotions. At a very high level, the image recognition process involves taking an image as an input (e.g., static frame of a CEO's face) and transforming the image into a field of weighted pixels to code eight facial emotions (*Anger*, *Contempt*, *Disgust*, *Fear*, *Happiness*, *Neutral*, *Sadness*, and *Surprise*) that have long been established as universal across cultures (Ekman & Friesen, 1971). The weights are generated by minimizing a loss/error function that compares the input image to images from a prior training set that has been coded for facial emotions.

To illustrate our methodology, we use an archive of video interviews with CEOs and founders conducted as part of Harvard Business School's “Creating Emerging Markets” project (publicly available for academic use in teaching and research). The archive consists of a collection of oral history transcripts—as well as their corresponding video recordings—of interviews with the CEOs of 69 unique organizations; the interviews were conducted from 2008 to 2018. CEOs came from a diverse set of countries, representing Asia, Africa, the Middle East, and Latin America. We used each of the CEO interview transcripts to code the variance of sentiment and topic entropy measures. We also coded the corresponding videographic material to generate facial expression scores for the eight emotions outlined earlier. Through a factor analysis, we then used our interview text sentiment scores (measures derived from the topic model of the interview texts) and our video-based facial expression sentiment variables to construct five distinct communication styles, which we label *Excitable*, *Stern*, *Dramatic*, *Rambling*, and *Melancholy*. Even within our highly selected sample of "star" emerging market CEOs, we find meaningful variation in CEO communication styles. We find that these five factors describe 87% of the variance among our video- and text-based variables.

For the purposes of illustration, we then engage in a proof-of-concept analysis to demonstrate the value of using our methods to synthesize text sentiments and facial expressions into CEO communication styles. The analysis employs a deductive approach in which we evaluate the suggestion from Helfat and Peteraf (2015) that firm leaders' language and oral communication are correlated with firm-level dynamic capabilities related to reconfiguration. Specifically, we collect data on acquisitions made by the CEOs' firms, a likely signal of asset reconfiguration. We conduct our analysis with both our full sample of CEOs and a subsample of 46 "active" CEOs, that is, those in our sample who were still performing the role of CEO at the time of their interviews. Our results reveal that CEOs who exhibit dramatic styles in their speech are less likely to oversee acquisitions. Our analysis also reveals the value of synthesizing textual sentiment and facial expressions into CEO communication styles, as opposed to analyzing firm-level outcomes using text and video data separately.

Finally, we compare our method to the prior methods of analyzing videographic data, and we outline the advantages of using our approach. As an example, we replicated the "videometric" method from prior work (Hill et al., 2019) by performing analyses using 100 human coders. Our analyses revealed a strong correlation between facial expressions coded by human coders and the state-of-the-art ML algorithm; however, as we argue later, the ML-based method has significant advantages in reducing research costs and time.

Our paper contributes to the literatures on managerial cognitive capabilities and CEO communication. By explicating our methods, we describe a new approach to operationalizing *communication styles* which, as yet, has been only invoked conceptually. We also demonstrate empirically the value of synthesizing multiple methods to generate CEO oral communication styles (that capture both verbal and nonverbal aspects of communication, as outlined by Helfat & Peteraf, 2015); and we compare our methods, which represent the state-of-the-art facial emotion recognition methods, to prior research methods. Most importantly, our methodological exposition opens up the possibility of strategy researchers embracing these methods and working with large repositories of textual, image, and video data across a variety of settings. Thus, the findings from our analysis are meant to illustrate the promise of our approach with ample room for future investigations.

2 | CEO COMMUNICATION: PRIOR THEORY AND METHODS

The CEO arguably occupies the most central and important leadership role at any firm, as he/she is principally charged with setting firm strategy (Hambrick & Mason, 1984). One of the most important ways that a CEO might influence firm strategy is by communicating his/her ideas to internal and external stakeholders (D'Aveni & MacMillan, 1990; Lefebvre, Mason, & Lefebvre, 1997; Yadav, Prabhu, & Chandy, 2007). In fact, Bandiera, Lemos, Prat, and Sadun (2013) and Bandiera et al.

(2018) measure how CEOs spend their time and show that a disproportionate percentage (85%) of CEO time is spent on activities that might involve communication (e.g., activities such as meetings, public speeches, phone calls, and conference calls).

From a theoretical standpoint in the strategy literature, CEO communication has been viewed as a core managerial cognitive capability that underpins the firm-level dynamic capability of reconfiguring. In the dynamic capabilities literature, reconfiguring is instrumental in achieving strategic asset alignment and overcoming resistance to change. As Helfat et al. (2007) argue, in the face of a change in the external environment, "reconfiguring" involves the acquisition of new assets, as well as the enhancement and/or reconfiguring of existing assets through innovation. Helfat and Peteraf (2015) establish a link between CEO communication and reconfiguring, outlining several characteristics of oral communication by CEOs and their effects on individual workers and firm strategy. The communication style of top managers in general, and the way in which they communicate a vision for the organization in particular, can inspire workers, encourage initiative, and drive entrepreneurial growth (Baum, Locke, & Kirkpatrick, 1998; Westley & Mintzberg, 1989). Managerial skill in using language, such as through impromptu talks, flow of words, and articulation in conversation, may affect worker response to change initiatives (Helfat & Peteraf, 2015, p. 843).

The authors also distinguish between "oral language" (i.e., what the CEOs say) and "nonverbal" communication (i.e., how they say it). In fact, Helfat and Peteraf (2015) argue that nonverbal behavior such as facial expressions and gestures can convey a range of information, including a person's opinions, values, cognitive states (such as comprehension or confusion), physical states (such as fatigue), and emotions. As the authors state, CEOs can use oral language and nonverbal communication "to facilitate strategic change within organizations and drive alignment by orienting members toward common goals" (Hill & Levenhagen, 1995; Helfat & Peteraf, 2015, p. 843). Together, these verbal and nonverbal forms of expression, in addition to written communication, constitute a CEO's *communication style*.

The empirical literature in strategy has long studied the effect of CEO communication on firm-level outcomes; however, the focus has been almost entirely on the content of written communication, rather than nonverbal and/or verbal forms of expression. Yadav et al. (2007) coded CEO communication using letters to shareholders that were featured in firms' annual reports. Using these data, the authors show that certain features of CEO communication—specifically having greater internal and external focus—can have a "positive and long-term impact on how firms detect, develop, and deploy new technologies over time" (Yadav et al., 2007, p. 84). Similarly, D'Aveni and MacMillan (1990) compared senior managers' letters to shareholders during demand-decline crises for 57 bankrupt firms and 57 matched survivors. The authors found that under environmental uncertainty, not only do surviving firm CEOs pay disproportionate attention to the output environment of the firm, but their communication to shareholders also more strongly reflects these structural differences in their attention. CEO communication also has been studied in the strategy literature on cognitive frames (the lenses that are shaped by their past experiences and through which CEOs interpret external stimuli). Kaplan (2008) uses CEO letters to shareholders and content analysis to measure managerial cognition.

In our review of strategy research on CEO communication, the workhorse methodological tool has been content analysis of CEO written communication, such as the analysis of CEO letters to stakeholders (Salancik & Meindl, 1984; Watzlawick et al., 1967). A recent literature looks at the text of earnings conference call presentations (Pan et al., 2018). The method of choice in the recent literature has been the LIWC method. As Gamache and McNamara (2019) explain, LIWC contains predesigned and pre-validated dictionaries of words measuring the positive and negative emotions within the text. LIWC is one of the several other text analysis tools used in the literature, having been increasingly adopted in strategic management research (Crilly, 2017; Kanze, Huang, Conley, & Higgins, 2018; Lungeanu, Paruchuri, & Tsai, 2018; Pan et al., 2018). Using this method, the authors find that negative media reactions to the announcement of a major acquisition is correlated with the degree to which the CEO and the firm engage in subsequent acquisition activity. The authors also find that the CEOs' "temporal focus," that is, the degree to which CEO attention is directed toward the past (coded using CEO letters to shareholders and employing the LIWC method) influences how sensitive CEOs are to media coverage. In another recent paper, Pan et al. (2018) use the LIWC method to code the level of "concreteness" in the top managers' language in earnings conference call presentations; they find that the use of concrete language by CEOs is correlated with positive investor reactions.

Only very recently have strategy scholars started paying attention to coding and studying CEO oral communication using videographic data.[1] Petrenko et al. (2016) developed a "videometric" method where third-party coders viewed snippets of CEO videos of varying lengths and rated each focal CEO on narcissism using a seven-point Likert scale. Using this videometric measure, the authors report a positive correlation between narcissism and measures of corporate social responsibility (CSR). More broadly, in a working paper, Hill et al. (2019, p. 2) define videometrics as a method that "uses third-party ratings of video samples to assess individuals' characteristics with psychometrically validated instruments of the measures of interest." In the accounting literature, scholars have used similar methods to code how CEOs' visual characteristics correlate to firm outcomes. Blankespoor et al. (2017) use 30-s content-filtered video clips of initial public offering (IPO) roadshow presentations to develop a measure of "investor perception" of a CEO and find that this measure is positively related to pricing at all stages of the IPO. The authors employ 900 workers on Amazon's Mechanical Turk (MTurk) to code 224 thin slices of videos created from the roadshow presentations and ask MTurk workers to use a seven-point Likert scale to provide their perceptions about a CEO's competence, trustworthiness, and attractiveness after watching the CEO's roadshow presentation.[2]

Despite the burgeoning interest in utilizing videographic data, the approaches reviewed above have been limited to analyzing snippets of videos, given the constraints of human coding. Furthermore, the content of what is being communicated has been overlooked in the stand-alone videometric analysis. However, recent advances in ML techniques now present strategy scholars with a chance to push the methodological boundaries to study CEO oral communication further. In fact, given that these ML algorithms could be applied to both text and facial image data, we now have an opportunity to begin to deliver on what Helfat and Peteraf (2015) have argued for: the systematic study of both verbal and nonverbal CEO oral communication. We now outline our methods, dataset, and results.

[1]Arguably, an important reason why researchers have neglected CEO oral communication has been the absence of methodologies hitherto that can perform "unsupervised" analysis using large datasets of CEO oral communication. In fact, Kaplan (2008) acknowledges this constraint and justifies the use of CEO written communication in her analyses by saying, "other kinds of statements by CEOs, such as those obtained through interviews or surveys, might initially appear to be attractive (data) sources, but they are impractical for larger samples of firms over long periods" (Kaplan, 2008, p. 679).

[2]In the broader management field, there is a related literature of coding still images of CEO faces and linking the coded measures to firm and individual performance. Graham, Harvey, and Puri (2016) study the facial traits of CEOs using nearly 2,000 subjects and link facial characteristics to both CEO compensation and performance. In one experiment, the authors use pairs of photographs and find that subjects rate CEO faces as appearing more "competent" than non-CEO faces. Halford and Hsu (2014) employ a sample of photographs for S&P 500 CEOs and find that facial attractiveness of the CEOs, coded from the still photographs, is positively correlated with firm returns. ten Brinke and Adams (2015) code facial expressions of CEOs from still photographs and find that when the face of the CEO exhibits happiness while he/she is tendering an apology following some firm transgression, the negative returns of the firm are heightened.

3 | A NEW APPROACH TO MEASURING COMMUNICATION STYLES

3.1 | Overview

We develop an approach that synthesizes methods for coding unstructured text data from oral communication and video data from the corresponding speakers to measure communication styles. We do so in the spirit of Helfat and Peteraf (2015, p. 837), who identify verbal and nonverbal "oral language" as one of the chief inputs to a manager's cognitive capabilities; it can have a profound influence on strategic decision making. Communication styles have been analyzed in a variety of settings, from physician-to-patient interactions (Buller & Buller, 1987) to political speeches (Perloff, 2008). Although a variety of definitions exist, Norton's (1978, p. 99) is arguably the most generalizable across contexts: "The way one *verbally* and *paraverbally* interacts to signal how literal meaning should be taken, interpreted, filtered, or understood" (emphasis added).

We illustrate a general approach for coding CEO communication style, with two major aims for strategy researchers. First, because our method brings together verbal text data and nonverbal facial expression data, we demonstrate how researchers can better understand how these two dimensions of communication interact to capture the unedited expressions of organizational leaders. Strategy researchers have begun to draw on ML methods to analyze and categorize large corpora of text data (e.g., Menon et al., 2018). However, in the context of CEO communication, text data are codified and written, representing the *edited* thoughts and views of their authors. Intent, attitudes, and views are often conveyed in nonverbal expression which, in some cases, can bring nuance to our understanding of content that is spoken and, in other cases, can contribute to surfacing a speaker's "authentic" perspective on a given matter. Therefore, our approach allows researchers to gain insight into *unedited* attitudes and feelings directly from the speaker.

Second, we *synthesize* our methods in a way that can be generalized across different sources of text and video data. Specifically, we bring attention to the widespread availability of video data. Although the video data for our sample of CEOs comes from a curated online archive of video interviews, we remind researchers that large online platforms (such as YouTube) and more focused news outlets (such as TechCrunch or CNN) contain searchable video archives of CEO speeches, interviews, and other forms of communication. Given that we increasingly consume information through online videos, researchers of CEO communication should be sensitive to online video platforms as an as yet untapped data source of managerial communication. Automated transcription software and facial expression coding algorithms constitute a set of freely available tools for the analysis of such widely available data. We describe how we take advantage of these types of tools in a way that can apply to other sources of video and communication data.

Figure 1 reports a general roadmap summary of our methodological approach.

The first three tracks in the flowchart represent the independent variables, or the measures we introduce to analyze the text and video: the topic model measures, which model the content of the interview text (Track A in Figure 1); the text sentiment measures (Track B in Figure 1), which represent the positive and negative sentiment reflected in the text; and the facial emotions gleaned from the video (Track C in Figure 1). Once these measures have been calculated, we use factor analysis to identify five clusters of CEO communication styles based on the text and video measures (Track E in Figure 1). As an illustrative example, we then examine how these inductive styles relate to the incidence of M&A transactions, represented in Track D. We describe each of these steps in detail, along with examples from our data, in the next section.

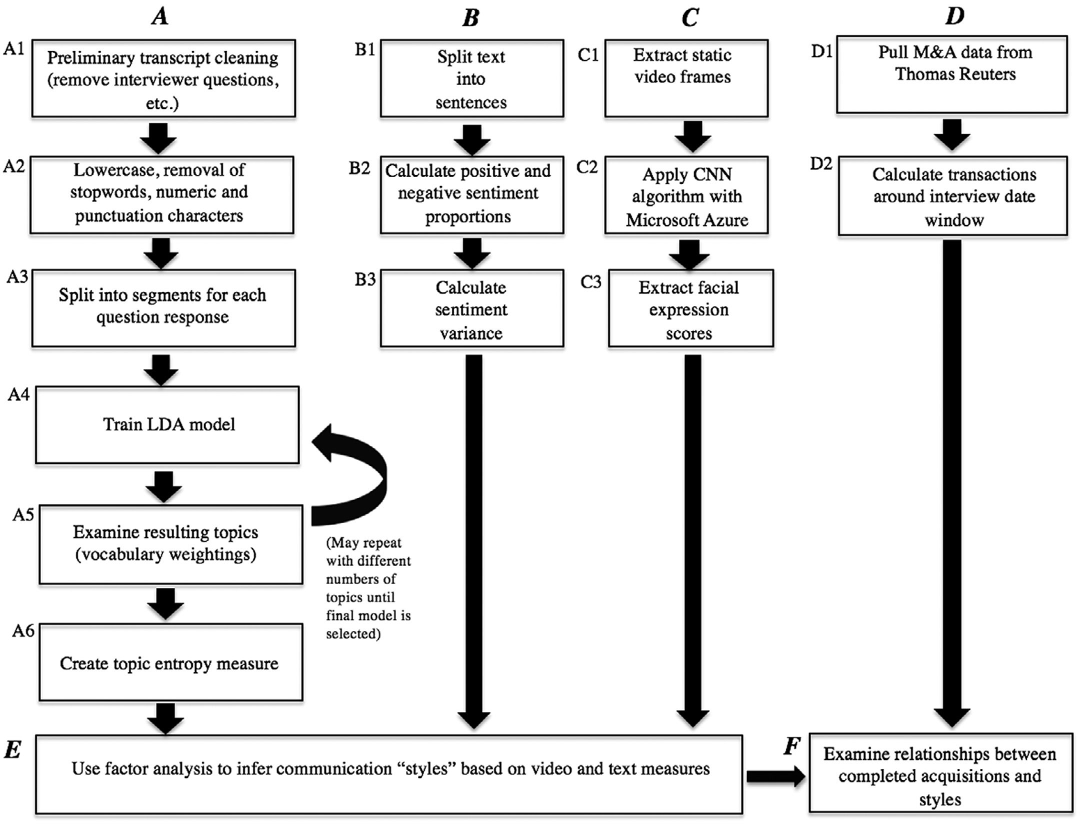

FIGURE 1 Overview of methodological approach

3.2 | Data and sample description

Our input text and video data come from an archive of video interviews with organizational leaders and founders conducted as part of Harvard Business School's "Creating Emerging Markets" project. The archive consists of a collection of oral history transcripts and video recordings of interviews with the leaders of 69 unique organizations; the interviews were conducted from 2013 to 2018 by researchers at the Harvard Business School. The individuals interviewed are typically entrepreneur founders, descendants of founders, or leaders. They may not be formally designated as "CEO," but are regarded as the leaders of companies or organizations regarded as iconic in their respective countries. The dataset, when we last accessed it (on October 15, 2018) had 115 interview transcripts but only 69 of these interviews had an accompanying video given that the video making process started in 2012. Given that our analysis requires the synthesis of text and facial image data, we based our analysis on the 69 interviews starting 2012, where we had both text and video data.

Given the unstructured nature of the interview format, the discussions give unique insight into each leader's unedited thoughts and attitudes but may not be reflective of CEO communication to other key stakeholders such as board members, stockholders, or the media. The interviews ranged from 1.5 to 2 hours in length and were transcribed and approved by the leaders prior to public distribution through the archive website. The dataset also has a key limitation: because our data come from interviews with "star CEOs from emerging markets," our sample is not representative of all organizational leaders. We position therefore our analysis as a proof-of-concept study that could be generalized to other CEO populations in future research. Also, the typical participant was over the age of 60. This arguably helped reduce informant bias, given that older informants could be more frank, as their words no longer affected their career prospects (Gao, Zuzul, Jones, & Khanna, 2017).

Examples of the organizations that the interviewed CEOs represented are the Tata Group from India, Aramex from the Middle East, and BTG Pactual from Brazil. We included 40 organizations from Asia, 12 from Latin America, nine from the Middle East and Turkey, and eight from Africa. The average interviewee was 68 years old at the time of the interview, and all video-recorded interviews occurred from 2012 to 2018.

3.3 | Coding interview text content: A ML approach

3.3.1 | Cleaning the text data

To code the verbal sentiment of CEO communication, we first obtained transcripts of the interviews with each of the 69 CEOs in our data. The interviews were, on average, 8,234 words in length, with a standard deviation of 3,458 words. A number of preprocessing steps were necessary prior to calculating our measures. In particular, we used only text that was spoken by the interviewee so that we did not simultaneously model the thoughts and opinions of the interviewer. Also, because several of the CEOs were interviewed in a language other than English (specifically, Spanish, Portuguese, or Turkish), we used the English translations of the interview transcripts as our input data. Admittedly, this might stand as a limitation of our approach, as our model might be accounting for a translator's own interpretations of a CEO's words rather than capturing the CEO's native tongue expressions.

The question-and-answer design of an oral interview provides a natural structure by which to segment each document. Specifically, after removing the interviewer questions, we treated each response as a separate "segment" for the LDA model (described below). We also calculated the average word length of segments for each CEO (*Average Answer Length*), considering that the tendency to respond to questions with brief, to-the-point answers or rather more long-winded replies may be a defining element in communication style. We again acknowledge a key limitation of the dataset: our data relies on CEO oral communication during semi-structured interviews with academics to conduct a proof-of-concept analysis; it is conceivable that the CEO might engage in different styles of communication with other key stakeholders.

3.3.2 | Content coding with the LDA topic model

Our next step was to use an unsupervised topic modeling approach to code the content of the CEOs' verbal responses. By unsupervised, we mean that the only input provided by the researchers in the model estimation is the overall number of topics. This approach allows the text to "speak for itself" in that the topics that emerge from the model are not influenced by what semantic subject matter the researchers might expect to find.[3] Given that the Latent Dirichlet Allocation method has been described in detail in prior strategy research (e.g., Kaplan & Vakili, 2015), we only briefly summarize how LDA generates topics from a set of documents.

The LDA model treats each document as a bag of words, meaning that the word order is not considered, and assumes an underlying random generative process in the creation of the "corpus"—or the set of documents being analyzed. It assumes that the collection of documents was generated by an imaginary probabilistic process, word by word, by first sampling a topic from a given document's

[3]This feature of the LDA method is in sharp contrast to one of the limitations of dictionary-based methods. As an example, Loughran and McDonald (2011) show that word lists developed for other disciplines "misclassify" common words in financial text.

distribution of topics and then sampling a word from that topic's word distribution. The sampling algorithm takes in the cleaned documents and then works backward, returning the most probable set of topics to have produced the given set of documents, if they had indeed been created in this imaginary way. A researcher can then infer the meaningful subjects represented by these topics and calculate the proportions of each document estimated to belong to each topic. We estimate the model on our sample of interview transcripts using the *topicmodels* package in R (Grun & Hornik, 2011).

Our final topic model for the set of transcripts contains 100 topics. This topic number was selected by triangulating across several different measures of model fit using the *ldatuning* package (please see our Supporting Information for a description of how we settled on 100 topics). The top 10 terms for each of the 100 topics can be viewed in Figure A3 in Appendix S1. The resulting model gives an intuitive summary of the semantic subjects discussed in the body of interviews. Some of the topics appear to be industry specific, while others are more general. Topic 75, for example, clearly seems to refer to marketing and branding ("brand," "brands," "market," "consumer," "strong") while Topic 22 seems related to retail ("stores," "store," "retail," "sell," "concept"). The more general topics appear to span both work-related subjects (Topic 38, which seems to pertain to corporate boards, e.g., "board," "members," "executive," "directors," "holding"), as well as personal subjects (Topic 55, seemingly about family history: "father," "brother," "grandfather," "died," "brothers"). This breadth reflects the variety of subjects encountered in the freewheeling interview format and the manner in which it provides a unique view into the thoughts of each CEO.

To generate document-level covariates from the topic model, we calculate the proportion of words belonging to each topic in each segment. As our corpus structure consists of long documents split into segments, we collapse each topic proportion back to the original document—that is, the interview transcript—by weighting by the length of each segment. The resulting covariates each have a value between zero and one, and the proportions of the 100 topics sum to one for each interview.

3.3.3 | Constructing measures from the topic model: *Topic Entropy*

A central component of communication style is the tendency to stay on subject versus a penchant for bouncing between different semantic topics (Čech, Garabík, & Altmann, 2015; Wang & Liu, 2018). We use the LDA topic proportions to capture this tendency by calculating a Shannon entropy measure for each CEO. The measure—calculated as $-\Sigma(p_i \times \log_2 p_i)$ in which each p_i represents the proportion allocated to each topic—captures the extent to which a given CEO's interview reflects attention to many different topics or is concentrated around just a few. Specifically, low values of entropy (those closer to 0) represent concentrated attention to few topics and high values represent attention to a diversity of topics. In short, entropy measures the breadth of information content in a body of text. We are careful not to assign a strict interpretation to what it means to exhibit high topic entropy as it could reflect an individual's tendency to bounce around different topics or a capacity to forge connections among diverse topics. Therefore, as a baseline, we suggest that higher values of entropy for an interview transcript signals a wider range of ideas and opinions, which communicates whether a CEO takes a more specialized or general approach to reflecting on personal and professional matters.

3.4 | Coding interview text sentiment: A dictionary approach

Sentiment analysis is an umbrella term referring to methods that measure the emotional valence of a document—that is, the extent to which a text expresses positive or negative sentiment. These methods are usually dictionary based. The sentiment measures in this paper are calculated using the *Syuzhet* R package (Jockers, 2017), which uses crowdsourced lexicons (Mohammad & Turney, 2013). The NRC lexicons used here correspond to two sentiment categories, positive and negative.[4] The salience of "sentiment" in organizational communication has been long theorized in the strategy and organizations literature; Neilsen and Rao (1987) view the dominant coalition in the organization (comprising the CEO) as "producers of meaning" and other organizational members as "consumers of meaning" with their own attributions regarding the organization, the motives of the elite, and their own needs, and sentiments. Recent literature has used measurements of sentiment to analyze how positive and negative coverage of firm behaviors affect managerial decisions (Pan et al., 2018; Shipilov, Greve, & Rowley, 2019) or to inform measures of CEO personality, which has implications for firm performance and strategic change (Harrison, Thurgood, Boivie, & Pfarrer, 2019).

Each term in the lexicon is categorized as having either a positive or negative sentiment. For example, the word "abandon" is assigned as a negative sentiment, while the word "ability" is given a positive sentiment assignment. This approach is somewhat crude, as it does not consider the context or word order of a phrase; but on balance, it typically performs nearly as well as more complex approaches (Mohammad, Kiritchenko, & Zhu, 2013). We sum the terms associated with each of the sentiments at the sentence level and then calculate the proportion of each document dedicated to each sentiment so that the values sum to one. This means that the measures of negative and positive text sentiment for a document are perfectly collinear. Therefore, in our generation of communication styles, we use only the measure of *Negative Text Sentiment*.

As an example of how this process works, consider the following passage from Anu Aga's (then Chairperson of Thermax, one of India's largest energy companies) transcript:

> I don't think joining HR was difficult, but what was difficult was getting back to work after a gap of many years. I wondered how I could be away the whole day and come home late, leaving the children without me. I kept thinking, what if my children or my mother-in-law got sick and needed me? I was a bit anxious about how the other professionals in HR who had studied HR would accept me. But I must say, we had a wonderful team.

This segment would be scored with five negative words ("difficult," "gap," "late," "sick," "anxious") and two positive words ("mother," "wonderful"). If this short section was the entire interview, these sentiment values would then be converted to proportions, with a value of 0.71 for *negative* and 0.29 for *positive* (the sentiment values will always sum to one). The higher negative value reflects that the segment dwells mainly on negative sentiment (concern and anxiety about returning to work) punctuated with some positive sentiment (warm thoughts about the team).[5] Across the entire interview, the sentiment values provide a picture of the extent to which the CEO prefers to reflect on negative emotions or adopts a more positive tone—a key component of style.

Beyond the average sentiment valence reflected in these measures, an additional style component might be the tendency to vary between positive and negative verbal sentiment over the course of a conversation, an aspect of emotional expressivity (Kring, Smith, & Neale, 1994) which may affect

[4]In future work, it may be useful to develop custom dictionaries specific to the purposes of strategy researchers.

[5]We acknowledge, however, that whether a CEO exhibits positive or negative sentiment in their interviews might also be related to variation in the settings in which they are interviewed. Our models capture and control for the time frames during which CEOs are interviewed, but not necessarily other elements of the interview setting. We also note that based on communication with the originators of the data collection, the settings in which the interviews took place do not vary meaningfully that we should expect there to be a major effect of the interview environment on what each interviewee says.

perceptions of leadership (Slepian & Carr, 2019; Van Kleef et al., 2009). To capture this, we calculated the *Text Sentiment Variance* across the different segments for each interview. This is measured by calculating the negative text sentiment for each question response and then taking the standard deviation of these values. Higher values of text sentiment variance will reflect an inclination to swing more widely between positive and negative language over the course of the interview.

3.5 | Coding videographic facial expressions: A ML approach

The third analytical tool employed in this paper uses supervised ML technology that takes a static facial image as input and generates as output, weights along eight facial expressions: *Anger*, *Contempt*, *Disgust*, *Fear*, *Happiness*, *Neutral*, *Sadness*, and *Surprise*. Ekman and Friesen (1971) first proposed that the human face could express seven basic emotions that persisted across world cultures—anger, disgust, fear, happiness, sadness, surprise, and contempt. An eighth category—neutral—is frequently used to describe the absence of emotional facial expression in the automated coding of facial expressions. The tool we use—the Microsoft Azure Computer Vision REST Application Program Interface (API)—was developed by Microsoft and builds on research by Yu and Zhang (2015) and for static frame, generates weights on these eight facial expressions as part of the standard output.[6] We first describe the algorithm underlying this tool and then explain the use of the technology in detail.

3.5.1 | The convolutional neural network algorithm

The API utilizes a version of a class of algorithms known as convolutional neural networks (Yu & Zhang, 2015). Arguably, this method is state-of-the-art and an area of active research in computer science (e.g., multimodal analysis in Chen et al., 2017). The technical details of the algorithm are vital for researchers of artificial intelligence, but for strategy researchers, we summarize the conceptual ideas.

A supervised neural network algorithm is implemented in three steps. In the first step, researchers employ a "training set" (frame-by-frame snapshots of a video, in our case) that is labeled according to the speaker's facial emotions. In the second step, the actual input image is transformed into a field of "weighted pixels" by using a neural network. These pixel weights are used to generate values for parameters such as "openness of the mouth," "curvature of lips," "dimples on the cheek," etc. These parameters are then used to generate "output values" for facial expressions of the input image. In the third step, the weights (on the pixels) are optimized based on minimizing a loss function/error function, where the error is coded based on the difference of the "output values" of facial expressions coded in the prior step and the "target value" of the same facial expressions. The target values of facial expressions are generated based on the same parameters used to generate the output values; however, unlike the "output values," the "target values" are based on data from the training set.

To explain how the neural network algorithm works, we build on the rich literature in the field of computer science of employing neural net (NN) methods. According to Duffner (2008), NN algorithms are inspired by the human brain and its capacity to perform complex tasks by means of interconnected neurons, each performing a relatively simple operation. Similar to the human brain, an NN is a trainable structure consisting of a set of interconnected "neurons," each implementing a very simple function. Collectively, the NN performs a complex classification function or an approximation task.

[6]Available at https://azure.microsoft.com/en-us/services/cognitive-services/emotion/.

In the case of facial image recognition algorithms, each "neuron" corresponds to a pixel in the image data. The task of the algorithm is to "read" the input image file and generate a set of "weights" to be assigned to pixels to code the parameters of interest, such as "skin color" or "openness of the mouth." As an example, skin color might indicate the existence of eyes or hair (vs. rest of the face). An open mouth with dimples on the cheeks might indicate the facial expression of "happiness."

In a simple, brute-force approach, the NN algorithm could consider each pixel of the input image data and assign weights to every pixel to compute the parameters of interest and minimize the loss function (as described earlier). However, this would be a case of "over-specification" and would be computationally intractable for most image datasets (Dietterich, 1995; Mullainathan & Spiess, 2017). Instead, the NN algorithm conducts localized optimization, where pixels in a "neighborhood" are assigned weights to successively generate higher level weights. A convolutional NN algorithm builds on this principle by converting an input image into a multilayer hierarchical structure where the first layer relates to the input image, the next few layers relate to "shallow" collections of pixels, pixels are grouped based on their neighborhood (e.g., neighborhoods comprising the edges of the image, part of the nose, part of the eye, etc.), and the subsequent layers relate to "deep" explorations of distant neighborhoods covering the entire face. To summarize, the weights are iteratively chosen to minimize the loss function described earlier. Once the final weights are assigned, the algorithm generates scores for the facial expression emotions.

3.5.2 | Capturing and coding static frames from videos

We use the Microsoft API, which outputs facial expression scores for a set of images with the NN algorithm described above. Before using the Microsoft tool, researchers must prepare the facial image data; if the facial image data is available as part of a video file (as in our case), this entails capturing individual static image frames from the video file. This task can be achieved by using media player applications. We use a cross-platform, open-source "VLC media player," which allows for the capture and export of static image frames from video data by using its "scene video filter" option. Settings within the VLC filter preferences can be used to adjust the number of frames extracted and their associated file names and file types. We captured one static image frame per second of video footage and, most importantly, used only the static image frames that related to the face of the CEO. In other words, we dropped from the sample all static image frames related to the face of the interviewer and static image frames without any facial images (e.g., title frame).

It is important to note that the algorithm implemented by the Microsoft API is also able to detect a face amid other objects in a static image. Although seemingly an obvious and necessary feature, this capability represents a major breakthrough in artificial intelligence image recognition technology, which makes the use of facial expression tools much more accessible across image and video data sources. In effect, the ability to detect faces removes the major barrier of having to manually crop images to make faces more prominent, a laborious pre-processing step.

Once the static images are ready for use, researchers can employ the Microsoft tool to generate facial expression scores. To do so, they must first apply for an API key from Microsoft Cognitive Services for permission to use the Face API. A free trial and set of API keys for the Face API is available to researchers through the Microsoft Cognitive Services website.[7] Signing up for the Face API grants a single user a key that permits processing up to 30,000 static images at a rate of 20 images per minute. The API returns emotion scores for the eight facial emotions, where each emotion receives a score between zero and one, according to the algorithm developed by Microsoft. That

[7]https://azure.microsoft.com/en-us/services/cognitive-services/face/.

data is reported back in a JSON file. We used SAS to collate the facial emotion scores and the frame number from the collection of JSON files.

3.5.3 | Facial expression data output

The data returned from the Face API assigns scores between 0 and 1 for each of eight different emotions—*Anger*, *Contempt*, *Disgust*, *Fear*, *Happiness*, *Neutral*, *Sadness*, and *Surprise*—for each image. The sum of the eight scores (for the eight emotions) for a given image is equal to 2. Therefore, a score for a given emotion can be interpreted as an indicator of the intensity of the emotion expressed relative to the other emotions that could be expressed. Because a set of images for a given interviewee represents one-second snapshots of the interviewee's video, taking the average score of *Fear*, for example, for the entire video gives a summary of the extent to which the individual on camera expressed fear. Figure 2 displays examples of static frames with the emotions recognized by the algorithm as having the highest scores. To assess the validity of the algorithm, in our Supporting Information, we summarize an analysis in which we compared the Face API-coded expressions to human-coded expressions for a selected set of facial images from our video data. The evidence shows that although the human coders in our sample do not align perfectly with the Face API's classification of facial expressions, there is considerable overlap. As a result, we take these results to indicate that we can treat the facial expressions coded by the Microsoft Face API with reasonable validity.

3.6 | Synthesizing interview content, interview sentiment, and facial expressions

3.6.1 | Discovering communication styles through factor analysis

In the next step, we use the facial expression and text sentiment scores along with measures from our topic model to discover communication styles among the CEOs in our sample (Section E of Figure 1). The reasoning behind this step is that facial cues and spoken content are likely to reveal some information about the speaker's preferred mode of communicating. We are assuming in this analysis that the sentiments expressed in the speaker's words and facial expressions will help approximate communication styles in a way that the content of the interview cannot. For example, our approach approximates how a bystander's perception of a CEO could be serious and buttoned-up or loose and informal, after witnessing the CEO conduct an interview. We use factor analysis to discover the styles in our sample although any form of multi-dimensional scaling could serve a similar purpose.

3.6.2 | Factor analysis results

In our factor analysis, we include 12 variables: the net negative text sentiment measure (*Negative Text Sentiment*), the text sentiment variance measure (*Text Sentiment Variance*), the average word length of each response (*Average Answer Length*), the topic entropy measure (*Topic Entropy*), and the eight facial emotion measures (*Anger*, *Contempt*, *Disgust*, *Fear*, *Happiness*, *Neutral*, *Sadness*, and *Surprise*). Employing the Kaiser-Guttman rule (retaining factors with an eigenvalue greater than 1), we obtain five factors constituting five different "styles."[8] We termed these five styles *Excitable*, *Stern*, *Dramatic*, *Rambling*, and *Melancholy* after examining the factor loadings, which are displayed in Table 1.

[8]For an implementation of a parallel analysis and a longer discussion of the number of factors retained, please see Figure A4 in Appendix S1.

FIGURE 2 Examples of static frames representing the eight facial expressions. *Note*: Score for corresponding emotion coded by Microsoft Face Application Program Interface displayed in bold text

The first factor, *Excitable*, is defined by consistently positive language (negative loadings on both negative text sentiment and text sentiment variance), as well as an association with fearful, surprised, and happy facial expressions. Those strong in this factor display significantly fewer neutral facial expressions, which is notable, as neutral is the most dominant facial emotion overall. The second factor, which we call *Stern*, is characterized by more angry, contemptuous, and disgusted facial expressions and less facial happiness; however, this factor is also associated with more neutral faces, leading us to interpret it as a stern, no-nonsense style.

We term the third factor *Dramatic*, as its strongest associations are with such disparate facial expressions as anger, disgust, happiness, sadness, and a lack of neutral expressions. We name the fourth factor *Rambling*, as it is most strongly characterized by long answer responses with high topic entropy. This style is also associated with facial contempt and happiness. Finally, we name the fifth factor *Melancholy*. This style loads heavily on facial sadness and contempt and negatively on facial happiness and anger.

In checking the stability of the style factors, we tested whether the same or different factors are revealed when only the text-based measures—or conversely, only the video-based measures—were used. On their own, the video-based measures produce two factors that appear similar to the ultimate *Excitable* and *Melancholy* styles. The text-based measures, when used alone in the factor analysis, produce only one factor that loads heavily on the *Average Answer Length* and *Topic Entropy* measures and negatively on the *Text Sentiment Variance* measure. This factor corresponds most closely to our *Rambling* style, though this style was also characterized strongly by facial contempt and happiness, features that are not evident from the text-based factor analysis alone. In general, the observed styles do

TABLE 1 Factor loadings

	Factors, labeled by authors				
	Factor 1	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**
Variable	Excitable	Stern	Dramatic	Rambling	Melancholy
Negative text sentiment	−0.165	−0.078	−0.004	0.021	0.062
Text sentiment variance	−0.257	−0.169	0.055	−0.186	0.086
Average answer length	−0.001	0.174	−0.227	0.406	0.115
Topic entropy	0.067	−0.022	−0.187	0.478	0.171
Video: Anger	0.093	0.710	0.358	−0.069	−0.456
Video: Contempt	0.072	0.386	0.223	0.618	0.486
Video: Disgust	0.082	0.727	0.499	0.009	−0.043
Video: Fear	0.758	0.267	−0.333	−0.281	0.083
Video: Happiness	0.341	−0.723	0.385	0.325	−0.303
Video: Neutral	−0.835	0.308	−0.418	−0.129	0.098
Video: Sadness	0.293	−0.132	0.411	−0.457	0.677
Video: Surprise	0.708	0.187	−0.583	0.019	−0.064

Note: Displayed are the loadings of the component variables on the first five factors of the factor analysis, with given names for the factors. Loadings greater than 0.3 are highlighted in a lighter shade, and loadings less than −0.3 are highlighted in a darker shade. The dataset, when we last accessed it (on October 15, 2018) had 115 interview transcripts but only 69 of these interviews had an accompanying video given that the video making process started in 2012 and between 2008 and a large part of 2012, only audio interviews were conducted. Given that our analysis requires the synthesis of text and facial image data, we based our analysis on the 69 interviews, where we had both text and video data. The factors remain similar if only the sample of 46 active CEOs used in the correlational acquisitions analysis is used for the factor analysis.

not seem to emerge fully without the inclusion of both text- and video-based measures, supporting the idea that both verbal and paraverbal communication are crucial to establishing a communication style.

4 | CEO COMMUNICATION STYLES AND FIRM OUTCOMES: A PROOF-OF-CONCEPT ANALYSIS

4.1 | M&A Activity and CEO communication

To illustrate the value of synthesizing multiple methods and measures to generate CEO communication styles, we perform proof-of-concept analyses and correlate the styles to a firm-level outcome. Our choice of the firm-level outcome is driven by prior literature. Helfat and Peteraf (2015) state that managers' cognitive capabilities for language and communication are likely related to dynamic managerial capabilities for reconfiguration. In turn, as Helfat et al. (2007) argue, reconfiguration often involves “asset orchestration,” that is, the selection, modification, configuration, and alignment of tangible and intangible assets. In fact, prior literature suggests that reconfiguration and asset orchestration might involve mergers and acquisitions (Capron, Dussauge, & Mitchell, 1998; Helfat & Peteraf, 2003; Teece, 2007). Given this, we choose acquisitions as the firm-level outcome variable for our analyses.

4.1.1 | Firm outcome measures

Acquisition data was compiled using the SDC Platinum database from Thomson Reuters. Restricting our data to all U.S. and non-U.S. targets within the date range 1980 to present, we accessed every M&A transaction in which the companies run by the interviewees acted as the acquiring company, during the tenure of the CEO.[9] Once that list of transactions had been compiled, we restricted our data to the time window surrounding the interview, calculating the number of completed acquisitions in the one-, three-, and five-year windows before and after the interview date. In our subsequent analysis, we restricted our data to the set of interviewees who were acting CEOs at the time of the interview. The number of completed acquisitions ranges from zero to six, with mean values of 0.22, 0.39, and 0.59 transactions within each of the respective time windows.

4.1.2 | Relationships between styles and completed acquisitions

Table 2 examines the relationships between the CEOs' scores on the style factors and the number of completed acquisitions within 1, 3, and 5 years of the interview (Columns 1–3, respectively). The models estimate OLS regressions following the specification:

$$Acquisitions_{\mathrm{i}} = \beta_0 + \beta_1\, Excitable_{\mathrm{i}} + \beta_2\, Stern_{\mathrm{i}} + \beta_3\, Dramatic_{\mathrm{i}} + \beta_4\, Rambling_{\mathrm{i}} + \beta_5\, Melancholy_{\mathrm{i}} + \beta \boldsymbol{X}_{\mathrm{i}} + \varepsilon_{\mathrm{i}}$$

The covariate vector X_{i} includes gender and region indicators for Asia, Africa, and Latin America (the omitted region being the Middle East), as well as fixed effects for each year in the sample (2012–2018).

The coefficients for each style variable can be interpreted as the number of additional (or fewer) completed acquisitions associated with a one standard deviation increase in a given style's factor score. We observe, for example, that a one standard deviation increase in the *Dramatic* style factor is associated with 0.26 fewer acquisitions within 1 year of the interview (Column 1, $p = .14$). This effect increases to 0.39 fewer acquisitions within 3 years ($p = .04$) and 0.56 fewer acquisitions within 5 years ($p = .06$). The manner in which the effect size increases with time might point to a cumulative "CEO effect"; while a one-year time frame might be subject to some noise, as the time window increases, the impact of the CEO could rise (similarly, the effect size associated with female CEOs grows over time).

Why might there be a relation between CEOs exhibiting a "dramatic" style and CEOs pursuing fewer M&A transactions? Are they less growth oriented or perhaps more likely to pursue different growth strategies? Questions such as these might be tested with larger, more robust samples (where researchers can control for fixed firm effects, endogeneity of the choice of CEO, etc.). We employ this correlational analysis as an illustrative example of how these methods may be used in inductive theory building. While this analysis is correlational in nature and intended primarily as a proof-of-concept, we encourage readers to use it as a springboard for related work on CEO communication and strategic decision making.

4.1.3 | Illustrating the value of synthesis

Why use the styles gleaned from the factor analysis, as opposed to the individual measures produced by the sentiment and facial analyses? Table 3 explores how the component measures

[9]Systematic data on divestitures, which also constitute asset reconfiguration, were not readily available.

TABLE 2 Estimated coefficients from linear regression of number of acquisitions on CEO communication style

	DV: Completed acquisition within *x* years of interview		
	x = 1 year Model 1	*x* = 3 years Model 2	*x* = 5 years Model 3
Excitable	0.04 (0.12)	0.18 (0.16)	0.52 (0.41)
	$p = 0.74$	$p = 0.27$	$p = 0.20$
Stern	−0.05 (0.06)	−0.06 (0.10)	−0.05 (0.15)
	$p = .44$	$p = .57$	$p = .74$
Dramatic	−0.26 (0.18)	−0.39 (0.19)	−0.56 (0.29)
	$p = .14$	$p = .04$	$p = .06$
Rambling	0.05 (0.09)	0.08 (0.12)	0.15 (0.18)
	$p = .58$	$p = .49$	$p = .41$
Melancholy	−0.04 (0.07)	−0.08 (0.10)	−0.10 (0.13)
	$p = .62$	$p = .40$	$p = .45$
Year FE	Yes	Yes	Yes
Region FE	Yes	Yes	Yes
Gender	Yes	Yes	Yes
Observations	46	46	46
Adjusted R^2	.33	.32	.39

Note: Each cell displays the OLS estimated coefficient with robust SE in parentheses and p-value underneath. Models control for gender, region, and year of interview. Of the 69 videos in our sample, this analysis utilizes a subsample of interviews related to 46 "active" CEOs, that is, those in our sample who were still performing the role of CEO at the time of their interviews.
Abbreviations: CEO, chief executive officer; DV, dependent variable; FE, fixed effects; OLS, ordinary least squares; SE, standard errors.

perform relative to the synthesized styles. Using OLS regressions with the number of completed acquisitions as a dependent variable, we compare the explanatory power of the text sentiment measures (Column 1), the answer length and topic entropy measures (Column 2), the facial expression scores (Column 3), all component measures together (Column 4), and the synthesized style scores (Column 5), in turn.

The text sentiment measure has very little explanatory power on its own, with an adjusted R^2 of −.01. The segment length and topic entropy measures perform only slightly better, with an adjusted R^2 of .04. The video measures explain the most variance on their own, but are difficult to interpret, with only one measure showing a meaningful relationship with the outcome (*Fear*, $p = .06$).[10] Comparing Columns 4 and 5 demonstrates why the clustered styles are useful: while the R^2 drops slightly, the adjusted R^2 rises from .00 to .09. Furthermore, the negative relationship between the firm acquisition activity and the *Dramatic* style emerges in a way that would not

[10]Given that "fear" and "sadness" load together on the "Dramatic" style with the lowest *p*-values, it is possible that their simultaneous inclusion is redundant. It is also possible that "fear" and "sadness" are largely what explain most of the variance in our facial expression measures. To assess this possibility, we conducted our analysis by creating factors without measures of "fear" and "sadness" to test whether such factors would also contribute to explaining acquisitions. We find that without "fear" and "sadness," the factors that emerge from the other emotions still contribute considerably to predicting acquisitions, demonstrating the value of synthesizing facial expression and textual data to measure communication style. These additional analyses are available from the authors upon request.

TABLE 3 Linear regression models demonstrating value of synthesis

	DV: Completed acquisitions within 5 years of interview				
	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**
Negative text sentiment	−3.72 (4.24)			−2.09 (4.25)	
	$p = .38$			$p = .63$	
Text sentiment variance	−1.04 (4.83)			4.25 (8.61)	
	$p = .83$			$p = .63$	
Average answer length		0.01 (0.01)		0.01 (0.01)	
		$p = .26$		$p = .32$	
Topic entropy		0.65 (0.97)		1.17 (1.11)	
		$p = .51$		$p = .29$	
Video: Anger			−7.26 (5.11)	−7.47 (5.77)	
			$p = .16$	$p = .20$	
Video: Contempt			−4.35 (13.79)	−11.38 (20.77)	
			$p = .76$	$p = .59$	
Video: Disgust			−5.40 (19.79)	−8.11 (22.02)	
			$p = .79$	$p = .72$	
Video: Fear			118.28 (61.69)	123.47 (61.48)	
			$p = .06$	$p = .05$	
Video: Happiness			1.26 (1.83)	1.12 (1.94)	
			$p = .50$	$p = .57$	
Video: Sadness			−6.27 (3.80)	−6.01 (3.62)	
			$p = .10$	$p = .10$	
Video: Surprise			2.23 (7.19)	1.48 (7.35)	
			$p = .76$	$p = .84$	
Excitable					0.52 (0.41)
					$p = .20$
Stern					−0.05 (0.15)
					$p = .74$
Dramatic					−0.56 (0.29)
					$p = .06$
Rambling					0.15 (0.18)
					$p = .41$
Melancholy					−0.10 (0.13)
					$p = .45$
Year FE	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes
Gender	Yes	Yes	Yes	Yes	Yes
Observations	46	46	46	46	46

TABLE 3 (Continued)

	DV: Completed acquisitions within 5 years of interview				
	Model 1	Model 2	Model 3	Model 4	Model 5
R^2	.26	.30	.40	.47	.39
Adjusted R^2	−.01	.04	.03	−.0001	.08

Note: Each cell displays the OLS estimated coefficient with robust SE in parentheses and p-value underneath. Models control for gender, region, and year of interview.
Abbreviations: DV, dependent variable; FE, fixed effects; OLS, ordinary least squares.

have been apparent from the component measures. The synthesized styles, therefore, provide us with a more intuitive and interpretable result than the component measures, with little loss of information.

5 | COMPARISON TO EXTANT TEXT-BASED AND VIDEOMETRIC METHODS

In our analysis of the interview transcripts, we employ both unsupervised topic modeling and more traditional dictionary-based methods to code sentiment. We also created two measures to reflect both the variance in sentiment expressed and the diversity of topics discussed, both of which we view to be relevant features of a communication style. Our approach is meant to illustrate the value of synthesizing these different elements of communication as a way of quantifying a leader's communication style.

Combining these features allows for greater flexibility than more standard approaches to textual analysis, such as LIWC, which rely on prevalidated dictionary measures. Most notably, LDA inductively models the content of a body of texts in a way that is not influenced by a researcher's priors and is difficult to achieve through set dictionaries (unless those dictionaries have been specifically designed for an empirical context). The LDA model estimated in this paper, for example, reveals topics about constructs of interest to strategy scholars (such as investment and corporate social responsibility), as well as topics highly specific to our data (such as Topic 64, which is about the tea industry). In addition, our new *topic entropy* measure provides a standardized method for estimating topic concentration that may be used across a variety of contexts, without being reliant on an established list of terms. This is not to say that dictionary-based methods are not valuable; we employ the crowdsourced *Syuzhet* lexicons in our sentiment analysis precisely because this remains a reliable and efficient way to measure sentiment. However, in our approach of combining both supervised and unsupervised methods, we aim to provide a model for a more flexible way of characterizing text-based communication that is less beholden to context.

Our approach also contributes to growing interest in the use of videometric methods to code visual expressions from videographic data that is now widely accessible (Blankespoor et al., 2017; Cade, Koonce, & Mendoza, 2018; Hill et al., 2019). Nevertheless, there are several challenges associated with implementing the videometric methods deployed in existing work. As such, we describe how our approach improves upon existing methods in terms of *efficiency*, *replicability*, and *adaptability*. As a baseline, we compare our method to the approach pioneered by Hill et al. (2019), but we raise similar examples from other recent work, as well.

Hill et al. (2019) describe the procedures involved in training a set of "raters" for a videometric coding task for a set of recorded interviews with CEOs (as described earlier). Although, they do not report the total length of time required to code their CEO videos from the start of rater recruitment to the final robustness checks, we can speculate that the task itself required a minimum of several days and could take as long as several weeks.

In terms of efficiency, our approach in using the Microsoft REST API for accomplishing the same videometric coding task eliminates the need to train human coders, given that it is a software tool accessible to any researcher with an internet connection. In addition, the processing time is dramatically shorter. Whereas it would be almost inconceivable to ask human coders to code every second of a video that is even an hour long, the REST API takes still frames of a video as input data that can then be processed in mere seconds, even for a video that is hours in length. Finally, because human coders are not required, researchers do not face the administrative cost of securing a physical environment for the coding task. Together, these efficiency advantages make videometric coding via an automated API like ours far more preferable because they lower the barriers to measurement for video data.

In addition, because our approach relies on software, researchers can follow our methods and reproduce the exact same measures with the same data. This is vital for efforts to replicate studies that use videometric data. Specifically, if one attempted to replicate a study that relied on human coding of video recorded facial expressions, it would be difficult to rule out that any unexpected differences in results might simply be a function of using different coders, for example. Our method facilitates the process of replication, making it easier for future researchers to confidently extend the results obtained using the Face API platform as well as verify the reproducibility of the analysis.

Finally, although by default, the Face API outputs ratings for a fixed set of eight different emotions, the algorithm also yields precise measurements of an individual's physical facial attributes, which might be adapted for other measurement purposes. For example, the algorithm measures the physical coordinates of one's eyes, nose, ears, and mouth, the presence of eye and lip makeup, the angle of one's head tilt, and the precise shape of one's head. These precise measurements serve as variables that researchers could potentially use to train new ML algorithms for identifying other expressions that might not be captured by the set of emotions we exploit in our analysis.

6 | DISCUSSION AND CONCLUSION

6.1 | Summary

We outline a novel synthesis of three methodologies—topic modeling (using unsupervised ML), sentiment analysis of text, and a cutting-edge facial image expression recognition (using supervised ML)—with an application to CEO oral communication. Exposition of these methodologies allows us to respond to the call made by Helfat and Peteraf (2015) to study verbal language and nonverbal communication, important inputs to managerial cognitive capabilities. Specifically, we collected and processed verbal and nonverbal data from a set of video-recorded interviews with CEOs and founders from emerging markets to show that both components are informative when identifying different communication styles.

6.2 | Contributions

Our analysis offers a glimpse of a potentially important predictor—CEO communication style—of firm behavior. Because our video data come from recorded sessions of unstructured interviews in which CEOs engage in free association with minimal prompting from an interviewer, our setting provides a unique perspective on how CEOs view what is important to them. Researchers have recently taken an interest in measuring how CEOs allocate their attention as a key input into understanding how they make decisions about firm strategy. In particular, researchers gather data by collecting information on how CEOs spend their days through detailed diaries (Bandiera et al., 2018). Our approach sheds light on several promising ways to further the literature on CEO oral communication by: (a) introducing how the synthesis of text- and video-based measures can generate CEO communication styles, (b) introducing a new measure of communication entropy (Shannon entropy) using the analysis of topics embedded in CEO communication text, and (c) revealing how communication style might be related to firm-level outcomes.

In the context of research on CEO communication, the set of methodologies developed in this paper could be used in the literature of cognitive frames (Kaplan, 2008), interpretation (Barr, 1998), and how CEOs spend their time (Bandiera et al., 2013, 2018). More broadly, these new methods for coding verbal and nonverbal communication could be particularly instrumental in research that analyzes natural language. As Suddaby and Greenwood (2005) outline, drawing from Burke's (1966) notion of language as “symbolic action,” several streams of research related to strategy employ the analysis of language. Important subfields of related research include semiotics (Barley, 1983), hermeneutics (Phillips & Brown, 1993), discursive analysis (Kilduff, 1993), narrative analysis (Boje, 1995), and rhetorical analysis (Freedman & Medway, 1994). Scholars in each of these subfields could benefit from using the methodologies outlined in this paper.

Related, we argue that the use of videographic data is necessary to measure the CEO's *communication style*. A communication style includes how one *verbally*—that is, *what* we say—and *paraverbally*—that is, *how* we say it—interacts to signal how what one says should be interpreted. By developing insight into what constitutes a CEO's communication style, we build directly on Helfat and Peteraf's (2015, p. 843) observation that “managerial skill in using language” can “inspire workers, encourage initiative, and drive entrepreneurial growth.” We argue that this “skill” can be captured and measured by synthesizing measures of verbal and nonverbal communication that are uniquely available through videographic data.

Our methodological exposition is relevant for the literature that uses language to assess how personality traits of CEOs relate to strategic change in the companies they manage (Chatterjee & Hambrick, 2007; Nadkarni & Chen, 2014). In a recent paper in this literature, Harrison et al. (2019) created personality measures using R's machine-learning capabilities in three stages: (a) text vectorization, (b) training and model selection, and (c) trait prediction. In the first stage, they used Word2Vec to extract language features from the larger text corpus of 3,573 CEOs. Our methods and exposition of coding CEO communication style could be helpful in further advancing this stream of research.

More broadly for strategy research in the age of Twitter, Instagram, and YouTube, these tools could be used by strategy scholars to code text, static images, and video data in a wide variety of settings. Arguably, a new set of methodologies to work with qualitative data such as text, static images, and video images provides an empirical breakthrough. In fact, as a recent *SMJ* editorial persuasively argues, studies using qualitative empirical methods have been instrumental in advancing the field of strategic management (Bettis, Gambardella, Helfat, & Mitchell, 2015). The article outlines several qualitative methods that have been used in strategy research, including qualitative comparative analysis (Ragin, 2014), first- and second-order analysis (Gioia, 2014), the case study method (Eisenhardt & Graebner, 2007), and rhetorical analysis (Suddaby, 2014).

The exposition of our novel set of methodologies to utilize video data adds to the relatively thin literature on the use of historical data in strategy research. In particular, Jones and Khanna (2006) outline two dimensions of historical data that make it difficult for use in broad strategy research—such data is often "qualitative" and "small sample." The authors then suggest methods that strategy scholars could use to analyze historical data and list methods related to Boolean algebra (Ragin, 2014), string analyses (Abbott, 2001), and computational models (O'Rourke and Williamson, O'rourke & Williamson, 1999). Oral history data—especially that accompanied by images or video—is arguably an underutilized data source for strategy research, and it often shares the qualitative and small sample properties outlined by Jones and Khanna (2006); our novel set of methodologies provides strategy scholars yet another empirical tool to use to further historical analysis in strategy research. In effect, we show how, even with a small sample of interviews (n = 69), our approach (through segmenting each interview transcript) allows for a meaningful and replicable quantitative analysis through topic modeling, sentiment analysis, and facial image recognition analysis.[11]

6.3 | Limitations

Our study has several limitations. Because our data are limited to interviews with CEOs of firms in emerging markets, we cannot generalize our results about CEOs' emotions and topical attention to those in other settings, such as developed markets. We encourage researchers to adopt our methods for future projects that might examine such a comparison. As stated earlier, our proof-of-concept analysis relies on CEO communication with academics via semi-structured interviews. It is plausible that CEO communication is different with other internal and external stakeholders. Also, our correlational analysis relating CEO communication styles to M&A outcomes is meant to be expositional and not intended to generate inductive insights. Ultimately, we hope researchers will find our method generative as means to develop new ways to incorporate diverse data sources into the study of strategic decision-making and leadership.

In addition, in terms of data limitations, as Kaplan (2008) states, the study of video interview data suffers from the risk of retrospective bias, as managers would likely adapt their memories of their views in prior years to subsequent outcomes. Oral history data might suffer from biases attributed to data generated from situational interviews, as identified in the psychology literature (Latham & Saari, 1984). Additionally, oral history data might be relevant in establishing links to firm outcomes only if the leader being interviewed is in an "active" managerial role during the time of the interview; this prompted us to utilize only the subsample of "active CEO" data in our baseline analysis. Also, we note that our analysis is constrained by the lack of repeated observations of the CEOs over time; our analysis would be improved with longitudinal videographic data.

As for other technical limitations, we can account for differences only in the region-of-origin for our CEO interviewees and the firms they represent. In other words, because our CEOs represent emerging markets, we caution readers that the results of our analysis of how communication styles are related to firm outcomes might not generalize to samples of CEOs from other regions. However, as a feature of the interview data collection, the CEOs' regions are also associated with whether the interviews were conducted in English. For instance, most CEOs from South American countries were

[11]In Appendix S1, we list selected oral history archives (mostly housed in university libraries) that contain a diverse array of interviews with business leaders, covering a wide range of industries, regions, and topics. One notable resource is Columbia University's Oral History Archive, which has been widely acknowledged as the largest searchable database of oral history records in the world, giving access not only to audio and video records of interviews with business executives, but also their accompanying transcripts.

interviewed in their native Spanish, which means our analysis could incorporate only the English translations of their interview transcripts.

Finally, although the format of the textual data we used in our analysis was especially well suited for generating topic models, text from other videographic data, such as lengthy addresses or speeches, might not be. In our data, we split each interview transcript into text segments, each of which represents an answer to a question posed by an interviewer. Therefore, each interview text segment is relatively cohesive in content and consistent in length. However, in other videographic data, the text associated with an individual's speech might not be segmented as cleanly. In such instances, the researcher must define the segmentation as part of a preprocessing step. How text is segmented might then affect the ultimate results of a topic model, which has implications for the set of communication styles that might be discovered.

6.4 | Future directions and applications

While the current study represents our efforts to advance our understanding of how ML methods could be used gainfully in strategy research focused on CEO oral communication, future efforts could augment our current study in several ways. Future research could explore whether we learn more from studying oral communication—or learn something different—than from studying written statements of CEOs. While we focus coding on text and facial expressions of oral communication, it would be possible to additionally use voice intonation and code yet one more dimension of oral communication. Also, it would be interesting to investigate the sensitivity of topic model results to translation effects. Additionally, although our approach utilized unsupervised LDA to estimate topic models, it is possible that a supervised approach could produce additional insights on topic estimates (Ramage, Hall, Nallapati, & Manning, 2009). A supervised approach would require researchers to read through a sample of transcripts and to associate certain words with predetermined topics, giving the topic model a fixed prior method for structuring the relationship between estimated topics. A supervised approach is encouraged when the language used in a corpus of documents has excessive jargon, such that relevant experts would be able to identify which specific and salient words should cohere as a topic. The language in our interviews does not necessarily reflect the excessive use of jargon, but it is possible that other oral business histories would exhibit higher proportions of industry-specific terminology. Additionally, while we use "bag of words" methods to construct the topics, it might also be interesting to study how the order of words correlates with sentiments expressed. Finally, future research might augment our textual sentiment analysis by creating and using a lexicon of words curated from papers published in the field of strategy to code sentiments expressed in the words spoken or written by CEOs.

Our approach can also be extended to other sources of videographic data that are becoming widely available. Media organizations and firms alike frequently post videos to open access video platforms such as YouTube, many of which contain recordings of executives speaking and interacting. These videos can reveal a great deal about a CEO's leadership approach through verbal and nonverbal patterns that have been unexplored as yet. Therefore, an opportunity exists to collect a data archive of CEO videos, which might then be categorized along a number of dimensions—such as recordings of leaders speaking in formal versus informal situations. Future researchers might adapt our approach to generate stylistic profiles of each CEO. Possible questions that could be explored include how CEOs' communication styles affect perceptions of them as leaders. In addition, researchers might use videographic data to predict executives' ascension into CEO positions. The tools we present in our analysis would facilitate research into these new questions.

In conclusion, from the perspective of strategy research on CEO oral communication, we document a replicable method to measure communication styles based on topic modeling of text, sentiment analysis of text, and a state-of-the-art facial image emotion recognition algorithm. We demonstrate a novel synthesis of the three methods to generate CEO oral communication styles that incorporate both verbal and nonverbal aspects of communication. We exploit an underutilized type of data for strategy research, that is, oral history text and video data to develop a proof-of-concept of using our method and provide evidence suggestive of how communication style correlates with the firm-level outcomes, such as completed acquisitions. This result speaks to the importance of studying *both* verbal and nonverbal language in relation to cognitive capabilities related to reconfiguration, highlighted by Helfat and Peteraf (2015). Most important, our set of methods—and the exposition of synthesizing these methods—opens the door for strategy scholars to use easily available yet underutilized text, image, and video data in a wide variety of settings.

ACKNOWLEDGEMENTS

The authors would like to thank Connie Helfat and participants at seminars at the China Europe International Business School, Duke University, and University of Toronto for comments on a prior draft. Patrick Clapp provided excellent research assistance on this project.

ORCID

Natalie A. Carlson https://orcid.org/0000-0001-7439-2682

REFERENCES

Abbott, A. (2001). *Time matters: On theory and method*. Chicago, IL: University of Chicago Press.

Antweiler, W., & Frank, M. Z. (2004). Is all that talk just noise? The information content of internet stock message boards. *Journal of Finance*, *59*(3), 1259–1294.

Arts, S., Cassiman, B., & Gomez, J. C. (2018). Text matching to measure patent similarity. *Strategic Management Journal*, *39*(1), 62–84.

Athey, S. (2019). The impact of machine learning on economics. In A. K. Agrawal, J. Gans, & A. Goldfarb (Eds.), *The economics of artificial intelligence: An agenda*. Chicago, IL: University of Chicago Press.

Bandiera, O., Guiso, L., Prat, A., & Sadun, R. (2018). What do CEOs do? Review of Financial Studies, Forthcoming.

Bandiera, O., Lemos, R., Prat, A., & Sadun, R. (2013). *Managing the family firm: Evidence from CEOs at work* (NBER Working Paper No. 19722).

Barley, S. R. (1983). Semiotics and the study of occupational and organizational cultures. *Administrative Science Quarterly*, *28*(3), 393–413.

Barr, P. S. (1998). Adapting to unfamiliar environmental events: A look at the evolution of interpretation and its role in strategic change. *Organization Science*, *9*(6), 644–669.

Baum, J. R., Locke, E. A., & Kirkpatrick, S. A. (1998). A longitudinal study of the relation of vision and vision communication to venture growth in entrepreneurial firms. *Journal of Applied Psychology*, *83*(1), 43–54.

Bettis, R. A., Gambardella, A., Helfat, C., & Mitchell, W. (2015). Qualitative empirical research in strategic management. *Strategic Management Journal*, *36*(5), 637–639.

Blankespoor, E., Hendricks, B. E., & Miller, G. S. (2017). Perceptions and price: Evidence from CEO presentations at IPO roadshows. *Journal of Accounting Research*, *55*(2), 275–327.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, *3*, 993–1022.

Boje, D. M. (1995). Stories of the storytelling organization: A postmodern analysis of Disney as "Tamara-land". *Academy of Management Journal*, *38*(4), 997–1035.

Buller, M. K., & Buller, D. B. (1987). Physicians' communication style and patient satisfaction. *Journal of Health and Social Behavior*, *28*, 375–388.

Burke, K. (1966). *Language as symbolic action: Essays on life, literature, and method*. Berkeley and Los Angeles, CA: Univ of California Press.

Cade, N., Koonce, L., & Mendoza, Kim (2018). *Using video to disclose forward-looking information: The effect of non-verbal cues on investors' judgments* (Working Paper).

Capron, L., Dussauge, P., & Mitchell, W. (1998). Resource redeployment following horizontal mergers and acquisitions in Europe and North America, 1988-1992. *Strategic Management Journal*, *19*(7), 631–661.

Čech, R., Garabík, R., & Altmann, G. (2015). Testing the thematic concentration of text. *Journal of Quantitative Linguistics*, *22*(3), 215–232.

Chatterjee, A., & Hambrick, D. C. (2007). It's all about me: Narcissistic chief executive officers and their effects on company strategy and performance. *Administrative Science Quarterly*, *52*(3), 351–386.

Chen, M., Wang, S., Liang, P. P., Baltrušaitis, T., Zadeh, A., & Morency, L.-P. (2017). *Multimodal sentiment analysis with word-level fusion and reinforcement learning*. Proceedings of the 19th ACM International Conference on Multimodal Interaction, 162–171.

Crilly, D. (2017). Time and space in strategy discourse: Implications for intertemporal choice. *Strategic Management Journal*, *38*(12), 2370–2389.

D'Aveni, R. A., & MacMillan, I. C. (1990). Crisis and the content of managerial communications: A study of the focus of attention of top managers in surviving and failing firms. *Administrative Science Quarterly*, *35*(4), 634–657.

Dietterich, T. (1995). Overfitting and undercomputing in machine learning. *ACM Computing Surveys*, *27*(3), 326–327.

Duffner, S. (2008). *Face image analysis with convolutional neural networks*. (Doctoral dissertation). University of Freiburg, Freiburg, Germany.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, *17*(2), 124.

Eisenhardt, K. M., & Graebner, M. E. (2007). Theory building from cases: Opportunities and challenges. *Academy of Management Journal*, *50*(1), 25–32.

Freedman, A., & Medway, P. (1994). Locating genre studies: Antecedents and prospects. In A. Freedman & P. Medway (Eds.), *Genre and the new rhetoric* (pp. 1–20). London, England: Taylor & Francis.

Gamache, D. L., & McNamara, G. (2019). Responding to bad press: How CEO temporal focus influences the sensitivity to negative media coverage of acquisitions. *Academy of Management Journal*, *62*(3), 918–943.

Gao, C., Zuzul, T., Jones, G., & Khanna, T. (2017). Overcoming institutional voids: A reputation-based view of long-run survival. *Strategic Management Journal*, *38*(11), 2147–2167.

Gavetti, G. (2012). PERSPECTIVE—Toward a behavioral theory of strategy. *Organization Science*, *23*(1), 267–285.

Gioia, D. (2014). A 1st-order/2nd-order qualitative approach to understanding strategic management. In Primer: Qualitative research in strategic management. *Strategic Management Journal*. Retrieved from https://www.strategicmanagement.net/smj/overview/submission/editorials-and-primers

Glaeser, E. L., Kominers, S. D., Luca, M., & Naik, N. (2018). Big data and big cities: The promises and limitations of improved measures of urban life. *Economic Inquiry*, *56*(1), 114–137.

Graham, J. R., Harvey, C. R., & Puri, M. (2016). A corporate beauty contest. *Management Science*, *63*(9), 3044–3056.

Grun, B., & Hornik, K. (2011). Topicmodels: An R package for fitting topic models. *Journal of Statistical Software*, *40*(13), 1–30.

Halford, J. T., & Hsu, S. H. (2014). Beauty is wealth: CEO appearance and shareholder value.

Hambrick, D. C., & Mason, P. A. (1984). Upper echelons: The organization as a reflection of its top managers. *Academy of Management Review*, *9*(2), 193–206.

Harrison, J. S., Thurgood, G. R., Boivie, S., & Pfarrer, M. D. (2019). Measuring CEO personality: Developing, validating, and testing a linguistic tool. *Strategic Management Journal*, *40*(8), 1316–1330.

Helfat, C. E., Finkelstein, S., Mitchell, W., Peteraf, M. A., Singh, H., Teece, D. J., & Winter, S. G. (2007). *Dynamic capabilities: Understanding strategic change in organizations*. Malden, MA: Blackwell Publishing.

Helfat, C. E., & Peteraf, M. A. (2003). The dynamic resource-based view: Capability lifecycles. *Strategic Management Journal*, *24*(S1), 997–1010.

Helfat, C. E., & Peteraf, M. A. (2015). Managerial cognitive capabilities and the microfoundations of dynamic capabilities. *Strategic Management Journal*, *36*(6), 831–850.

Hill, A., Petrenko, O., Ridge, J., & Aime, F. (2019). Videometric measurement of individual characteristics in difficult to access subject pools: Demonstrating with CEOs. In D. J. Ketchen, Jr. & D. D. Bergh (Eds.), *Research methodology in strategy and management* (pp. 39–61). Bingley, UK: Emerald Publishing.

Hill, R. C., & Levenhagen, M. (1995). Metaphors and mental models: Sensemaking and sensegiving in innovative and entrepreneurial activities. *Journal of Management*, *21*(6), 1057–1074.

Jockers, M. (2017). Package 'Syuzhet.' Retrieved from https://cran.r-project.org/web/packages/syuzhet

Jones, G., & Khanna, T. (2006). Bringing history (back) into international business. *Journal of International Business Studies*, *37*(4), 453–468.

Kanze, D., Huang, L., Conley, M. A., & Higgins, E. T. (2018). We ask men to win and women not to lose: Closing the gender gap in startup funding. *Academy of Management Journal*, *61*(2), 586–614.

Kaplan, S. (2008). Framing contests: Strategy making under uncertainty. *Organization Science*, *19*(5), 729–752.

Kaplan, S., & Vakili, K. (2015). The double-edged sword of recombination in breakthrough innovation. *Strategic Management Journal*, *36*(10), 1435–1457.

Kilduff, M. (1993). Deconstructing organizations. *Academy of Management Review*, *18*(1), 13–31.

Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., & Mullainathan, S. (2017). Human decisions and machine predictions. *Quarterly Journal of Economics*, *133*(1), 237–293.

Kring, A. M., Smith, D. A., & Neale, J. M. (1994). Individual differences in dispositional expressiveness: Development and validation of the emotional expressivity scale. *Journal of Personality and Social Psychology*, *66*(5), 934–949.

Latham, G. P., & Saari, L. M. (1984). Do people do what they say? Further studies on the situational interview. *Journal of Applied Psychology*, *69*(4), 569–573.

Lefebvre, L. A., Mason, R., & Lefebvre, E. (1997). The influence prism in SMEs: The power of CEOs' perceptions on technology policy and its organizational impacts. *Management Science*, *43*(6), 856–878.

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *Journal of Finance*, *66*(1), 35–65.

Lungeanu, R., Paruchuri, S., & Tsai, W. (2018). Stepping across for social approval: Ties to independent foundations' boards after financial restatement. *Strategic Management Journal*, *39*(4), 1163–1187.

Menon, A., Tabakovic, H., & Lee, C. (2018). *Using machine learning to predict high-impact general technologies (working paper)*. Philadelphia, PA: Wharton School.

Mohammad, S. M., Kiritchenko, S., & Zhu, X. (2013). NRC-Canada: Building the state-of-the-art in sentiment analysis of tweets. *arXiv:1308.6242*.

Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word–emotion association lexicon. *Computational Intelligence*, *29*(3), 436–465.

Mullainathan, S., & Spiess, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives*, *31*(2), 87–106.

Nadkarni, S., & Chen, J. (2014). Bridging yesterday, today, and tomorrow: CEO temporal focus, environmental dynamism, and rate of new product introduction. *Academy of Management Journal*, *57*(6), 1810–1833.

Neilsen, E. H., & Rao, M. H. (1987). The strategy-legitimacy nexus: A thick description. *Academy of Management Review*, *12*(3), 523–533.

Norton, R. W. (1978). Foundation of a communicator style construct. *Human Communication Research*, *4*(2), 99–112.

O'rourke, K. H., & Williamson, J. G. (1999). *Globalization and history: The evolution of a nineteenth-century Atlantic economy*. Cambridge, MA: MIT Press.

Pan, L., McNamara, G., Lee, J. J., Haleblian, J., & Devers, C. E. (2018). Give it to us straight (most of the time): Top managers' use of concrete language and its effect on investor reactions. *Strategic Management Journal*, *39*(8), 2204–2225.

Perloff, R. M. (2008). *Political communication: Politics, press, and public in America*. Boca Raton, FL: Routledge.

Petrenko, O. V., Aime, F., Ridge, J., & Hill, A. (2016). Corporate social responsibility or CEO narcissism? CSR motivations and organizational performance. *Strategic Management Journal*, *37*(2), 262–279.

Phillips, N., & Brown, J. L. (1993). Analyzing communication in and around organizations: A critical hermeneutic approach. *Academy of Management Journal*, *36*(6), 1547–1576.

Ragin, C. C. (2014). *The comparative method: Moving beyond qualitative and quantitative strategies*. Oakland, CA: University of California Press.

Ramage, D., Hall, D., Nallapati, R., & Manning, C. D. (2009). *Labeled LDA: A supervised topic model for credit attribution in multi-labeled corpora*. Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing, 248–256.

Riedl, C., Zanibbi, R., Hearst, M. A., Zhu, S., Menietti, M., Crusan, J., … Lakhani, K. R. (2016). Detecting figures and part labels in patents: Competition-based development of graphics recognition algorithms. *International Journal on Document Analysis and Recognition*, *19*(2), 155–172.

Salancik, G. R., & Meindl, J. R. (1984). Corporate attributions as strategic illusions of management control. *Administrative Science Quarterly*, *29*(2), 238–254.

Shipilov, A. V., Greve, H. R., & Rowley, T. J. (2019). Is all publicity good publicity? The impact of direct and indirect media pressure on the adoption of governance practices. *Strategic Management Journal* (Forthcoming).

Slepian, M. L., & Carr, E. W. (2019). Facial expressions of authenticity: Emotion variability increases judgments of trustworthiness and leadership. *Cognition*, *183*, 82–98.

Suddaby, R. (2014). Editor's comments: Why theory?

Suddaby, R., & Greenwood, R. (2005). Rhetorical strategies of legitimacy. *Administrative Science Quarterly*, *50*(1), 35–67.

Teece, D. J. (2007). Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance. *Strategic Management Journal*, *28*(13), 1319–1350.

ten Brinke, L., & Adams, G. S. (2015). Saving face? When emotion displays during public apologies mitigate damage to organizational performance. *Organizational Behavior and Human Decision Processes*, *130*, 1–12.

Van Kleef, G. A., Homan, A. C., Beersma, B., Van Knippenberg, D., Van Knippenberg, B., & Damen, F. (2009). Searing sentiment or cold calculation? The effects of leader emotional displays on team performance depend on follower epistemic motivation. *Academy of Management Journal*, *52*(3), 562–580.

Wang, Y., & Liu, H. (2018). Is Trump always rambling like a fourth-grade student? An analysis of stylistic features of Donald Trump's political discourse during the 2016 election. *Discourse & Society*, *29*(3), 299–323.

Watzlawick, P., Bavelas, J. B., & Jackson, D. D. (1967). *Pragmatics of human communication: A study of interactional patterns, pathologies, and paradoxes*. New York, NY: W. W. Norton & Co.

Westley, F., & Mintzberg, H. (1989). Visionary leadership and strategic management. *Strategic Management Journal*, *10*(S1), 17–32.

Yadav, M. S., Prabhu, J. C., & Chandy, R. K. (2007). Managing the future: CEO attention and innovation outcomes. *Journal of Marketing*, *71*(4), 84–101.

Yu, Z., & Zhang, C. (2015). *Image based static facial expression recognition with multiple deep network learning*. Proceedings of the 2015 ACM on International Conference on Multimodal Interaction, 435–442.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Choudhury P, Wang D, Carlson NA, Khanna T. Machine learning approaches to facial and text analysis: Discovering CEO oral communication styles. *Strat Mgmt J*. 2019;40:1705–1732. https://doi.org/10.1002/smj.3067