Extracting organizational culture from text: the development and validation of a theory-driven tool for digital data

Michael Schachner, M. Murat Ardag, Peter Holtz, Johannes Großer, Carina Hartz, Hester van Herk, Michael Bender, Klaus Boehnke & Henrik Dobewall

To cite this article: Michael Schachner, M. Murat Ardag, Peter Holtz, Johannes Großer, Carina Hartz, Hester van Herk, Michael Bender, Klaus Boehnke & Henrik Dobewall (17 Jun 2024): Extracting organizational culture from text: the development and validation of a theory-driven tool for digital data, European Journal of Work and Organizational Psychology, DOI: 10.1080/1359432X.2024.2360225

To link to this article: https://doi.org/10.1080/1359432X.2024.2360225
Extracting organizational culture from text: the development and validation of a theory-driven tool for digital data

Michael Schachner a, M. Murat Ardaq b, Peter Holtz a, Johannes Großer b, Carina Hartz a, Hester van Herk a, Michael Bender a, Klaus Boehnke b and Henrik Dobewall c

aDepartment of Marketing, Vrije Universiteit Amsterdam, Amsterdam, Netherlands; bBremen International Graduate School of Social Sciences, Constructor University, Bremen, Germany; ‘Knowledge Construction Lab, Leibniz-Institut für Wissensmedien, Leibniz, Germany; ‡Department of Social Psychology, Tilburg University, Tilburg, Netherlands; ‡Faculty of Education, Research Unit of Psychology, University of Oulu, Oulu, Finland

ABSTRACT
Understanding organizational culture and practices (OCPs) is crucial for business success, yet current assessments rely heavily on self-reports. We developed the “Dictionary of Organizational Culture and Practices”, the first theory-based text analysis tool designed to quantify OCPs as an alternative to surveys. Our approach integrates computational natural language processing with expert human judgement, using a 26 million words large corpus consisting of employee reviews of 526 companies, 1,121 mission and vision statements, and 285 company websites. We replicate Hofstede’s six-dimensional model of OCPs, not to be confused with Hofstede’s widely known model of national culture. Criterion validity is established by comparing our tool to existing ones that used data-driven approaches. Notably, our tool demonstrates construct validity through the striking similarity of the dimensional structure with survey-based measures. The level of alignment varied depending on the type of text, with congruence coefficients exceeding .9. Our findings offer valuable insights for academics and managers seeking to understand and shape an organization’s culture more effectively.

Introduction

Culture matters for organizations. They seek to assess it, typically via self-report surveys (Hofstede et al., 1990; Maher, 2000; Schein, 2010). However, self-reports have limitations, and organizational culture is expressed in numerous other ways (Denison et al., 2014; Reader et al., 2020). Current and former employees anonymously review companies on workplace rating sites, and companies present themselves to the outside world on their own websites and in reports addressed to their stakeholders and the general public. Such digital data hold information about organizational culture that is fundamentally different from answers to questionnaires. They are rather unobtrusive indicators of organizational culture (Reader & Gillespie, 2022; Webb et al., 1966) – without the direct engagement of employees. We set out to assess whether models largely based on self-report survey data are applicable to this type of data and aim to create the first theory-driven text analysis tool for measuring organizational culture. We do so relying on the organizational-culture framework, the Multi-Focus Model on Organizational Culture by Hofstede et al. (1990), which is predominantly used in the business literature (Caprar et al., 2022).

The culture of an organization manifests in its organizational practices, and these are a good indicator of its functioning and adaptiveness. Organizational practices influence a variety of indicators, such as knowledge sharing and innovation promotion (Connelly et al., 2012), but also employee health (Varonen & Mattila, 2000). Consequently, organizational culture impacts a company’s performance. According to 92% of interviewed North American senior executives, the culture of a firm is a key determinant of its market value (Graham et al., 2016). Given these links, it is of obvious importance to assess organizational culture. Here, we develop and validate measurement tools for extracting organizational culture from text data.

There are countless definitions of culture (Baldwin et al., 2006), but it is most generally referred to as a shared meaning system (Fischer, 2009) or the collective programming of the human mind (Hofstede et al., 2010). Organizational culture, more specifically, can be defined as shared organizational practices that distinguish one organization from another. Practices are the observable and visible manifestations of organizational culture and are thus measurable (Hofstede et al., 1990).

Organizational culture research has overwhelmingly used survey data (Fischer et al., 2014; Hofstede et al., 1990; Maher, 2000; Schein, 2010). Structured self-reports are attractive (e.g., standardization, ease of administration; Cheung, 2021), but they are prone to common method variance, social desirability, and response style biases (Baumgartner et al., 2021; He et al., 2014). At least partially, these limitations could be overcome by alternative ways of assessment. There has recently been a strong call to utilize complementary methodologies (Field et al., 2021; Shenkar et al., 2020) and digital data, such as words and phrases used within and around an organization. Studies so far range from textual analyses of the core values...
and goals of companies (Sull et al., 2019) to establishing a culture dictionary via machine learning (Li et al., 2021) and coding pictures of offices for capturing the organizational culture (Byrne et al., 2022). However, most importantly, it is not clear whether theoretically established constructs will be structurally similar when measured with different data sources and types; for example, is the relationship between the Hofstede organizational culture and practices (OCP) dimensions of Customer Orientation and Organizational Effectiveness the same in text data as in survey data?

There are different ways of assessing organizational culture in natural language. In-depth interviews are time-intensive and not easily scaled, whilst text analysis takes time when done manually. Automated text analysis, in contrast, is fast and objective in terms of treating every text produced about an organization equally. Most importantly, natural language-based analyses can be closer to the actual behaviour and may capture real-world organizational practices better (Boyd & Pennebaker, 2017). Employees’ unstructured and spontaneous talking or writing about the companies they work for in private is a more natural source than structured survey items. If the text has been written by or for companies, such as companies’ marketing texts, the results will reflect the organization’s own perspective; that is how it wants to be seen, not the organizational culture as such. When texts are selected carefully, text analysis may overcome some of the biases associated with self-completion surveys.

Increasing digitalization has made textual data more available for research (Pollach, 2012), such as texts published by companies about themselves (e.g., websites, annual reports), publicly accessible employee reviews (e.g., on www.Glassdoor.com or www.Kununu.com), or social media postings about and by companies (e.g., X, Facebook). Dedicated software and programming packages are widely available to analyse these textual data using different approaches. Dictionary-based approaches count the frequency of words appearing in a text, allowing for quantitative analysis of qualitative data; researchers assign words to known or unknown categories. The dictionary approach is rooted in the “bag-of-words” assumption where the word order is irrelevant, and documents are treated as containers (a bag). The “bag-of-words” assumption can be followed in supervised models where the researcher knows both input and output, or unsupervised models where categories are created automatically (Banks et al., 2018). The dictionary approach can be considered a supervised model in this respect.

The quantitative computational linguistic analysis of text data is a new measurement method available to science and practice. Semantic analysis can predict and understand psychological constructs from textual data (Arnulf et al., 2014), with open-ended questions demonstrating equal or higher validity and reliability compared to closed-ended numerical rating scales (Kjell et al., 2019). Grunenberg et al. (2024) outperformed traditional recruiter judgements in assessing applicants’ Big Five personality traits based on CVs and short text responses. Ponizovskyi et al. (2020) found correlations between measured and self-reported scores for 7 out of 10 basic human values from Schwartz’ theory (Schwartz, 1992). All these studies focus on individual-level psychological constructs; there is limited adoption in organizational contexts (Banks et al., 2018; Hickman et al., 2022; Schmiedel et al., 2019) and a research gap in validating established theories through text. Yet, two projects stand out in the literature of analysing organizational culture through text. First, relying on 1.2 million employee reviews (ER), the MIT SMR/Glassdoor Culture 500 project (Sull et al., 2019) assesses nine cultural values (agility, collaboration, customer, diversity, execution, innovation, integrity, performance, respect); the nine values were selected based on the frequency of more than 60 distinct values that companies listed in value statements on their websites. ER from Glassdoor, an employee review site, were extracted and assigned to topics. The final Culture 500 dictionary contains over 20,000 words assigned to 90 topics, finally clustered into nine values. The project is data-driven, but the authors mention a partial conceptual overlap with the 6-dimensional Organizational Culture Profile by O’Reilly et al. (2014). In the second project, Li et al. (2021) start with the most frequently mentioned values of S&P 500 firms on their websites: innovation, integrity, quality, respect, and teamwork. They utilize the Q&A section of over 60,000 earning call transcripts of over 7,000 companies. While the Culture 500 project by Sull et al. (2019) relied on a “painstaking process” (p. 8) of a lot of manual reviewing and mapping terms to topics, Li et al. (2021) applied a computational linguistics approach with few human interventions. They apply the “word embedding model”, a model they use to identify words and phrases that commonly appear in the context of their five target values and their synonyms. The validation of their 1800-term dictionary with external indices suggests a correlation between their dictionary-based results and firms’ performance on the five indicators. Further, firms with high scores on these indicators demonstrate superior financial performance, particularly during economic hardship.

These two projects suggest that a computational linguistics approach generally works for assessing organizational culture from digital data; however, neither has there been objective evidence of validation published for the Sull et al. (2019) instrument, nor can the Li et al. (2021) instrument directly be linked to theory. Both instruments, respectively, rely on only one type of text – the Culture 500 dictionary is created and applied to employee reviews and is not even expected to work with other types of text; Li et al.’s dictionary was tested only on earning calls, and it is not clear how the dictionary would perform on other types of text.

**The present study**

Employing natural language processing techniques, we establish a dictionary building upon: Hofstede et al. (1990) organizational culture scale, referred to as the Multi-Focus Model by practitioners and which is still a standard reference study (Maddux et al., 2021). The Multi-Focus Model on Organizational Culture consists of six dimensions:

D1 Effectiveness, Means- vs Goal-Orientation: In a means-oriented culture, emphasis is on work processes (the “how”), whereas in a goal-oriented culture, emphasis is on achieving specific internal goals or results (the “what”).

---

References:


**Sull, D. S., et al. (2019).* Culture 500:* Companies that Deliver Business Results. MIT Sloan Management Review.


---

For a comprehensive list of references and further readings, please consult the original document.
D2 Customer Orientation, Internal vs External: Internally-driven cultures prioritize business ethics and self-determined customer service standards, while externally-driven cultures stress meeting customer needs pragmatically and achieving results.

D3 Control, Easy-going vs Strict: An easy-going culture values improvisation and has a fluid structure, while a strict one places a premium on cost-consciousness, punctuality, and seriousness.

D4 Focus, Local vs Professional: In a local culture, employees identify with their boss or unit and tend to be short-term focused, while in a professional culture, they identify with their profession or job content.

D5 Approachability, Open vs Closed: Open cultures readily welcome newcomers and are inclusive, while closed cultures are less accessible and more protective of internal knowledge.

D6 Management Philosophy: Employee vs Work-Oriented: Employee-oriented organizations prioritize employee welfare, even at the expense of work, whereas work-oriented cultures emphasize task completion, potentially at the expense of employees.

Hofstede et al.’s organizational culture approach has been validated and applied frequently (Caprar et al., 2022; Verbeke et al., 2000).

The product of our study will be a text analysis tool, the “Dictionary of Organizational Culture and Practices” (DOCP). Our dictionary is based on the methodology developed by Ponizovskiy et al. (2020). They created a personal values dictionary using publicly available texts to measure human values according to Schwartz’s (1992) theory. Our dictionary quantifies organizational culture based on the Hofstede-based Multi-Focus Model. The entries of the dictionary are words and phrases, each being assigned to one or several dimension poles such as a goal-, customer-, or employee-oriented culture. Using natural language processing software such as LIWC (Linguistic Inquiry and Word Count; Tausczik & Pennebaker, 2010) or the corresponding R packages (Benoit et al., 2018), the DOCP might be used to analyse any text containing information about an organization’s culture and practices. We validate the dictionary against a dictionary on corporate cultural values of Li et al. (2021) because it is established in the literature of LIWC-like dictionaries on organizational cultural values and has been validated against a number of external measures such as operational efficiency or firm value, and is publicly accessible. We finally compare the data structure of dictionary-based results in text to the survey results of the well-established and widely used model originating from Hofstede et al. (1990).

Our study is the first, to our knowledge, with a theory-driven rather than a data-driven dictionary for assessing organizational culture, grounded in Hofstede’s et al. (1990) seminal work on measuring organizational culture through practices. In addition, unlike the previous studies by Li et al. (2021) or Sull et al. (2019) that have focused on a narrow selection of text, we use a wide range of textual types around organizations to create our dictionary of words and phrases used within and around organizations: mission and vision statements (MVS), company websites (CW), employee reviews (ER), interview transcripts, and terms used in various organizational practices surveys items (see Table S1 in the supplementary material for a more detailed comparison of studies).

Method

The Dictionary of Organizational Culture Practices (DOCP) was developed in English, following the conventional steps to build such a dictionary (Ponizovskiy et al., 2020; Tausczik & Pennebaker, 2010) as can be seen in Figure 1: (1) obtaining the source data (various text corpora), (2) generating and selecting candidate words, (3) mapping the candidate words to the dimensions to develop the provisional dictionary, (4) refinement, and (5) validation. Steps 1, 3, and 4 are mostly guided by human expert judgement, whereas Steps 2 and 5 apply data science algorithms and computational linguistics.

To address the limitation of single-sourced texts of earlier studies (Li et al., 2021; Sull et al., 2019), we chose text corpora in Step 1 that can be divided into three main categories: companies talking about themselves, employees talking about these companies, and generic texts describing organizational culture and practices (OCPs; for a detailed overview of sources for candidate words and their reasons for inclusion see Table S2 in the supplementary material). For companies talking about themselves, we obtained mission and vision statements (MVS) from 1,121 companies with more than 420,000 words and website texts of 25 million words from 285 companies overlapping with the companies for which we had MVS. For employees talking about these companies, we included 2.5 million words of reviews of 372,447 employees working in 257 matched companies that represented various industries. For generic texts describing OCPs, we included item texts of various questionnaires that assess organizational culture and interview questions from in-depth interviews on OCPs.

Candidate words and phrases are referred to as n-grams (see Table 1 for examples). Due to the very different sizes and nature of source data, different methods were used to generate candidate words in Step 2. The first method is based on frequency; the second method involves instrumentalizing a pre-trained natural language model (Honnibal et al., 2020) to extract noun phrases as n-grams. Companies with diverse organizational cultural profiles were selected based on high or low scores on the six dimensions of our theoretical model. These scores were derived from pre-existing survey assessments. Employee reviews and website texts were used as data sources. The supplementary material (S3) provides a detailed description of the data extraction process.

We ended up with a list of 17,522 terms in total. The authors went through the terms to categorize them into relevant keywords (for example, “professional development”, “strong support”, “lack motivation”) and non-relevant keywords to be eliminated (for example, “tea people”, “consulting firms”, “comfy shoes”). This elimination procedure gave us 3,134 keywords categorized as related to OCPs. An additional 514 words and phrases were manually selected from survey items of the instruments measuring OCPs and from interview questions used by the consulting company Hofstede Insights Ltd to assess OCPs. These 514 words and phrases were directly assigned to the poles of the six dimensions. For a detailed overview and references of the sources used, see Table S2; Table 2 provides an overview of the number of n-grams extracted from the various source data.
In **Step 3**, 11 certified consultants were recruited from within Hofstede Insights Ltd as experts to allocate n-grams to dimension poles. All experts had previously been trained in working with the Multi-Focus Model and hence had good working knowledge. The experts, who were financially compensated, mapped altogether 3,442 unique n-grams to the 12 poles of the 6 organizational-culture dimensions (Hofstede et al., 1990) of the Multi – focus Model or categorized the n-gram as not related to OCPs. To complete this procedure, we presented the n-grams to the experts as keywords in context (KWICs) – eight words before and after the keywords were displayed based on the context in which the n-gram appeared in the original text corpora (Manning & Schuetze, 1999). The online appendix (S4) contains the concrete instructions that were sent to the experts. As the selected n-grams appeared in the original text corpora in many different contexts, the same n-gram was displayed in two different contexts to every expert.

Two experts mapped each n-gram, resulting in up to four different contexts. The candidate words were allocated to one or more of the dimension poles of the Multi-Focus Model to link the terms to the existing theory. In total, 8,690 mappings were completed in 3,119 employee contexts and 5,570 company contexts; hence, on average, every expert completed 790 mappings. We subsequently coded the dimension allocations by the experts as shown in Table S5 to establish the provisional dictionary.

This preliminary version of the DOCP enabled us to create the document by term matrices for the different corpora. During the last stage of the development phase, **Step 4**, we relied on the subsets of three corpora: Mission and Value Statements, Employee Reviews, and Company Websites. We used only a randomly selected 70% of the corpora when conducting the following analyses, reserving the remaining 30% for estimating the internal consistency of the DOCP dimension poles.

During this stage, we further applied an exploratory factor analysis with weighted least squares extraction and oblique rotation to the raw word counts for each pole of the six dimensions, with the aim of n-gram reduction and verification. We eliminated the words and phrases with factor loadings lower than .10 to eliminate words and phrases that were unrelated to

<table>
<thead>
<tr>
<th>Table 1. Examples of n-grams, the building blocks of a dictionary.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type</strong></td>
</tr>
<tr>
<td>Uni-gram (1 word)</td>
</tr>
<tr>
<td>Bi-gram (2 words)</td>
</tr>
<tr>
<td>Tri-gram (3 words)</td>
</tr>
<tr>
<td>Four-gram (4 words)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2. Selection and distribution of n-grams across sources of data.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Source</strong></td>
</tr>
<tr>
<td>From mission and vision statements and website texts</td>
</tr>
<tr>
<td>From employee reviews</td>
</tr>
<tr>
<td>From survey items</td>
</tr>
<tr>
<td>From in-depth interviews and thick-description of company environment</td>
</tr>
</tbody>
</table>
other terms belonging to the same dimension pole or not distinctive enough to be kept in the refined dictionary. A larger reference value for the required minimum factor loading would have reduced the number of included words unnecessarily. Then, we checked for internal consistency of the DOCP dimension poles estimated using the binary (one-hot) scoring method: this method uses a binary count of the presence/absence of a word in a text and the reliability is calculated based on tetrachoric correlations of this binary data. This concludes the development phase.

As a first assessment of the structure, we visualized the refined DOCP dictionary to gain insight into the structure underlying the words and phrases mapped to the 12 dimension poles. We performed a homogeneity analysis through optimal scaling using the Homals package in R (de Leeuw & Mair, 2009).

In Step 5, we validated the refined dictionary. Except for the mission and vision statements (MVS), we used different text data sources for the validation: instead of publicly available employee reviews (ER), we used structured ER obtained from employee satisfaction surveys and provided by Great Place to Work.\(^2\) Company websites were replaced with publicly available annual reports (AR) of 124 stock-listed companies. As MVS were used as input for developing the dictionary and are thus prone to overfitting the data, the structured ER are similar in type to the publicly available ER, and AR are a type of text data that had not been used for the development of the dictionary altogether. Using different sets of data allowed us to gain insight into the range of applicability of the dictionary outside data sets used for development.

An overview of the validation steps and the data sources used in the respective analyses is provided in Figure 2.

To assess the criterion validity of our dictionary (Step 5a), we estimated the associations between DOCP dimension scores and the related constructs of Innovation, Integrity, Quality, Respect, and Teamwork (Li et al., 2021) in the same set of text corpora. When applying a dictionary through the LIWC software (Tausczik & Pennebaker, 2010), the percentage of n-grams in the texts that were allocated to a dimension or pole was calculated. Our dictionary measured each dimension pole separately. To get a meaningful dimension score, we subtracted the left dimension pole score from the right. A correlation between these two dictionaries is indicative of the criterion validity of the DOCP.

To assess construct validity (Step 5b), we tested the similarity of the data structure between survey data results and text data results. We created multidimensional scaling (MDS) solutions for each source of data (survey data, MVS, ER, AR) with the SMACOF package in R (Mair et al., 2022), and then Procrustes rotated the MDS solutions of the text corpora towards the survey data as our “ground truth” (Boyd et al., 2021). A strong similarity of the MDS solutions across types of data collections, that is, text versus self-report surveys, is indicative of some degree of construct validity of the developed dictionary.

The analyses were conducted in the statistical programming environment R (v3.6.3 and v4.3.0, respectively). The following packages/libraries were used: homals v1.0.10; LIWC-22 v1.3.0; psych v2.0.7; quantedv2.0.1; quantedatextstat v0.90; quantedatextmodels v0.90; smacof: 2.1.5; spacyr v1.2.1; Tableau v2024.1.0.

Results

The first version of the dictionary contained 3,305 n-grams with 1,705 cross-mappings. The agreement between raters in one
identical set of 500 KWICs mapped by two expert raters was 65%: Out of 100 words and phrases that were allocated to a dimension pole, 65 allocations were the same, which is considered a good number in practice. Overall, allocations were unequally distributed across the 12 poles – the goal orientation pole had the most allocations with 943 mappings, while the easy-going work discipline pole had the least with 245 mappings, showing that our expert raters found that the presented text reflects some dimensions more prominently than others. Table 3 exemplifies the allocation of n-grams to the dimensions and poles, and Table S6 shows an extract of the first dictionary entries in LIWC format. The refinement process reduced the number of n-grams to 2,870 and the cross-mappings to 1,671. The refined dictionary is available at the open science framework (https://osf.io/x5wqt/).

The internal consistencies of allocated keywords were high, between .81 and .95 across OCP dimensions and corpora (mission and vision statements, company websites, and employee reviews). Detailed internal consistency results are displayed in Table S7.

The poles of our dictionary are not independent as included n-grams loaded on more than one dimension pole. To assess the relationship between dimension poles based on the expert mappings, Figure 3 plots the results of a homogeneity analysis (explained variance: 26%; de Leeuw & Mair, 2009) using the dictionary as data source. The axes reveal a semantic space where proximities indicate thematic co-occurrence. The further away from the origin, the stronger the relations are. Positions on either side of the origin indicate a strong negative relationship.

The left side indicates a highly professional culture focused on progress and change, represented through keywords such as “ongoing change,” “employee morale,” “global network,” “best information,” “progression opportunities,” or “amazing team.” Within Hofstede et al. (1990) framework, the Multi-Focus Model, this type of culture is primarily included in the dimensions D1 Effectiveness (goal orientation: results are more important than processes) and D4 Focus (professional: high identification with the job), and to a lesser extent through D5 Approchability (open, regarding communication and diversity), and D6 Management Philosophy (employee orientation: taking into account personal problems of employees) as indicated by the arrows in Figure 3.

In contrast, the right side captures a competitive, dysfunctional culture with an unpleasant work environment, represented through keywords such as “unethical behaviour,” “ultra conservative approach,” “bad planning,” “centric management,” “poor work environment,” or “unshared information.” In line with theoretical expectations according to the Hofstede et al. (1990) framework, these features are captured by the opposing poles of the corresponding dimensions mentioned for the left side of the axis: a focus on means (D1 – processes more important than results), work (D6 – high pressure to perform tasks), a closed (D5 – untransparent and inaccessible) and local (D4 – loyalty more important than the content of the job) culture.

The top aligns with a business-oriented culture focused on clarity, represented through keywords such as “attention,” “consistent processes,” “clear objectives,” “global structure,” “clear rules,” or “highest standards.” These keywords are reflected in the dimensions D2 Customer Orientation and D3 Control within Hofstede et al. (1990) framework: interestingly, both poles of these dimensions belong to this section of the semantic space.

Finally, the bottom combines a people-oriented culture open to diversity, represented through keywords such as “support employees,” “developing people,” “supporting women,” “responsibility,” or “better work-life balance.” Notably, none of the Multi-Focus Model dimensions clearly capture this section, though employee orientation (D6 Management Philosophy) comes closest.

The spread of dimension poles based on the expert mappings differs from the expected distribution of the Multi-Focus Model dimensions as none of the theoretical dimensions by Hofstede et al. (1990) correspond to the lower part of the semantic space in Figure 3, and some dimension poles cluster together instead of being distributed throughout the space. Notably, the themes of the semantic space could equally well be allocated to other organizational culture models such as O’Reilly’s

Table 3. Example n-grams (uni-grams, bi-grams, tri-grams and four-grams) as they were allocated to the dimension poles.

<table>
<thead>
<tr>
<th>D1 – Effectiveness</th>
<th>D2 – Customer Orientation</th>
<th>D3 – Control</th>
<th>D4 – Focus</th>
<th>D5 – Approachability</th>
<th>D6 – Management Philosophy</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEANS-ORIENTED</td>
<td>INTERNALLY DRIVEN</td>
<td>EASY-GOING WORK DISCIPLINE</td>
<td>LOCAL</td>
<td>OPEN SYSTEM</td>
<td>EMPLOYEE-ORIENTED</td>
</tr>
<tr>
<td>commitments</td>
<td>rewards</td>
<td>delegator</td>
<td>authority</td>
<td>diverse</td>
<td>compensation</td>
</tr>
<tr>
<td>competing</td>
<td>company priorities</td>
<td>flexible</td>
<td>corporate hierarchy</td>
<td>collaborative environment</td>
<td>employee friendly</td>
</tr>
<tr>
<td>objectives</td>
<td>rules and policies</td>
<td>schedules</td>
<td>rogue</td>
<td>nice working conditions</td>
<td>good worklife balance</td>
</tr>
<tr>
<td>outdated work</td>
<td>try new things</td>
<td>managers</td>
<td>primadonna</td>
<td></td>
<td></td>
</tr>
<tr>
<td>policies</td>
<td></td>
<td>senseless</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>short term</td>
<td>practice what they</td>
<td>safe to take risks</td>
<td>managers</td>
<td>positive welcoming work</td>
<td>local hard working</td>
</tr>
<tr>
<td>bandaid approach</td>
<td>preach</td>
<td></td>
<td>key performance indicators</td>
<td>environment</td>
<td>employees</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GOAL-ORIENTED</td>
<td>EXTERNALLY DRIVEN</td>
<td>STRICK WORK DISCIPLINE</td>
<td>PROFESSIONAL</td>
<td>CLOSED SYSTEM</td>
<td>WORK-ORIENTED</td>
</tr>
<tr>
<td>challenges</td>
<td>customers</td>
<td>deadlines</td>
<td>knowhow</td>
<td>criticize</td>
<td>investment</td>
</tr>
<tr>
<td>annual goals</td>
<td>client requirements</td>
<td>formal procedures</td>
<td>continual improvement</td>
<td>distrust</td>
<td>increasing workload</td>
</tr>
<tr>
<td>cutting edge technologies</td>
<td>highest quality service</td>
<td>more detailed information</td>
<td>highest quality products</td>
<td>fear driven environment</td>
<td>endless mandatory overtime</td>
</tr>
<tr>
<td>long term growth</td>
<td>winning in the</td>
<td>too many</td>
<td>long term action plans</td>
<td>corrupt senior leadership team</td>
<td>nonexistent work life balance</td>
</tr>
<tr>
<td>ability</td>
<td>marketplace</td>
<td>administrative tasks</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
organizational culture profile (adaptability, integrity, collaborative, results-oriented, customer-oriented, and detail-oriented; O’Reilly et al., 2014) or the three practice dimensions by Fischer et al (employee practices, innovation practices, and formalization practices; Fischer et al., 2014).

To assess criterion validity, we calculated the Pearson correlations (Table 4) between our DOCP dimensions (rows) and the Corporate Cultural Values (Li et al., 2021, columns) across the data sets selected for validation: mission and vision statements (MVS), employee reviews (ER), and annual reports (AR). D1 Effectiveness (means minus goal orientation) was positively correlated with Innovation (MVS, .132), Integrity (MVS, .143; ER, .413), Quality (MVS, .078; ER, .389), Respect (MVS, .091; ER, .157), and Teamwork (MVS .107, ER, .584); This indicates that typical characteristics of a goal-oriented organization are an entrepreneurial attitude, creativity, productivity, and trustworthiness. D2 Customer Orientation (internal minus external) was positively correlated with Innovation (ER, .444), Quality (MVS, .112; ER .187), and negatively correlated with Integrity (ER −.275; AR, −.252). The highest correlation was found between D1 Effectiveness and Teamwork in the ER dataset (.584). All correlations from the employee review dataset were in line with conceptual expectations, whereas in the other datasets, there were occasionally counterintuitive correlations with sometimes opposing signs.

To assess construct validity, we compared the configurations of dimensions within four data sources: survey data (our “ground truth”), employee reviews, mission and vision statements, and annual reports. The DOCP captured the text content to a varying degree: 22% of the mission and vision statement, 18% of both the employee review and 12% of annual report words were listed in our dictionary and could be attributed to OCPs. Multidimensional scaling with Procrustean rotation towards the survey data solution was used for this analysis. MDS stress values and fit measures for each solution are reported in Table S8.
Table 4. Pearson correlations between our DOCP dimensions (rows) and the corporate cultural values (columns) by Li et al. (2021) based on LIWC results from three types of text data. The sources of the DOCP dimensions data are Mission and Value Statements (MVS, 1,091 companies), Employee Reviews (ER, 269 companies; aggregated to company level) and Annual Reports (AR, 124 companies).

<table>
<thead>
<tr>
<th>DOCP dimensions</th>
<th>Data type</th>
<th>Corporate Cultural Values (Li et al., 2021)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Innovation</td>
</tr>
<tr>
<td>D1 Effectiveness - means minus goal</td>
<td>MVS</td>
<td>.132**</td>
</tr>
<tr>
<td></td>
<td>ER</td>
<td>.032</td>
</tr>
<tr>
<td></td>
<td>AR</td>
<td>-.086</td>
</tr>
<tr>
<td>D2 Customer Orientation - internal minus external</td>
<td>MVS</td>
<td>-.003</td>
</tr>
<tr>
<td></td>
<td>ER</td>
<td>.444**</td>
</tr>
<tr>
<td></td>
<td>AR</td>
<td>.094</td>
</tr>
<tr>
<td>D3 Control - loose minus strict</td>
<td>MVS</td>
<td>-.081**</td>
</tr>
<tr>
<td></td>
<td>ER</td>
<td>.219**</td>
</tr>
<tr>
<td></td>
<td>AR</td>
<td>-.041</td>
</tr>
<tr>
<td>D4 Focus - local minus professional</td>
<td>MVS</td>
<td>.119**</td>
</tr>
<tr>
<td></td>
<td>ER</td>
<td>.449**</td>
</tr>
<tr>
<td></td>
<td>AR</td>
<td>.342**</td>
</tr>
<tr>
<td>D5 Approachability - open minus closed</td>
<td>MVS</td>
<td>-.096**</td>
</tr>
<tr>
<td></td>
<td>ER</td>
<td>-.341**</td>
</tr>
<tr>
<td></td>
<td>AR</td>
<td>-.036</td>
</tr>
<tr>
<td>D6 Management Philosophy - employee minus work</td>
<td>MVS</td>
<td>.085**</td>
</tr>
<tr>
<td></td>
<td>ER</td>
<td>-.386**</td>
</tr>
<tr>
<td></td>
<td>AR</td>
<td>.077</td>
</tr>
</tbody>
</table>

**Correlation is significant at the 0.01 level (2-tailed).
*Correlation is significant at the 0.05 level (2-tailed).

Figure 4. Three-dimensional MDS of survey data and text data (ER, MVS, AR) rotated towards survey data. The third dimension is visualized through the size of the dots. D1 Eff= Organizational Effectiveness; D2 Cust= Customer Orientation; D3 Cont= Control; D4 Foc= Focus; D5 Appr= Approachability; D6 Mgmt= Management Philosophy. Company representation: survey data 317 companies/150,604 respondents; Employee Reviews ER 206 companies, Mission and Vision Statements MVS 1,121 companies, Annual Reports AR 134 companies.

Survey Data (Figure 4, pane a): Six theoretical OCP dimensions were mapped in a three-dimensional space. The x-axis contrasted a goal-oriented and professional culture (left side) against a closed and work-oriented culture (right side). D3 Control and D2 Customer Orientation were central on the x-axis, opposing each other on the y-axis, and low on the third dimension, indicated through the size of the dots.

Employee Reviews (pane b) showed clear similarity to the survey data (congruence coefficient of .973). The main differences were the reversed positions of D1 Effectiveness and D4 Focus on the y-axis and the greater distance between D6 Management Philosophy and D5 Approachability, placing a goal-oriented (D1) and closed (D5) culture closer to a strict culture (D3 control) compared to the survey data.
Mission and Vision Statement Data (pane c) showed the biggest similarity to the survey data in terms of the congruence coefficient (.976). The dimensions were aligned as in the survey data, with the notable difference that, similar to the employee reviews, D6 Management Philosophy moved away from D5 Approachability compared to the survey data.

Annual Report Data (pane d) showed the least alignment with the survey data (congruence coefficient of .913). Yet, the six OCP dimensions were arranged similarly to the survey data, with the positions of D1 and D4 reversed on the y-axis, as in the employee review data.

All solutions resemble the results of the homogeneity analysis presented earlier (Figure 3), with the exception of D2 Customer Orientation manifesting on the opposing side of the y-axis. In summary, the configurations in different sources exhibited close alignment with the survey data, with some slight differences noted in the configurations of employee reviews, mission and vision statements, and annual reports. Congruence coefficients above .95 suggest equal configurations, while those above .85 are interpreted as fairly similar (Skimina et al., 2021). Notably, the distance between the closest dimensions in the survey data (D1 Effectiveness and D4 Focus; D5 Approachability and D6 Management philosophy) is greater in all three text-based configurations, suggesting that text analysis might be even better suited than questionnaires in capturing the conceptual differences between these dimensions.

Discussion

We created the Dictionary of Organizational Culture and Practices (DOCP) and applied it to different types of text with the aim to reproduce and extend the Multi-Focus Model on organizational culture (Hofstede et al., 1990) whose dimensions were originally established through the analysis of survey data. The approach adds to the literature and our understanding of organizational culture in several ways:

1) Adoption of a theory-driven approach: Unlike the conventional data-driven approach found in previous studies (Li et al., 2021; Sull et al., 2019), our study implements a theory-driven approach, reproducing an existing model: the Multi-Focus Model on organizational culture (Hofstede et al., 1990). The theory-driven approach is more challenging as the researcher needs to find theoretically defined dimensions from the text data rather than searching texts for frequently co-occurring n-grams and building a model around it. We succeeded in finding the a-priori defined OCP dimensions from text data and in showing that the dimensional structure of text data results is very similar to the dimensional structure of survey data results. Our tool provides a cost-effective and less biased method to compare organizational culture profiles using existing text data. It offers an alternative approach for future OCP studies based on the Multi-Focus Model.

2) Usage of diverse text data sources: Our study was also the first one, to our knowledge, that built an organizational culture dictionary from multiple sources of text data. We included text data from employee reviews, mission and vision statements, company websites, and various established survey instruments. Previous studies focused on one type of text data such as employee reviews (Sull et al., 2019) or earnings calls (Li et al., 2021) only. We showed that it is possible to develop a dictionary from multiple textual sources and to apply the developed dictionary to other types of texts (like annual reports) while maintaining the dimensional structure. However, based on the correlations with results of other, data driven tools (Table 4), results from different text sources cannot be interpreted in the same way. Consequently, a comparison of results from different text sources is likely not meaningful. This is an important finding that is likely relevant to other dictionary tools outside the organizational culture sphere.

3) Methodological contribution: Our approach was guided by predefined criteria for n-gram selection, refining the dictionary’s ability to differentiate between companies. We further demonstrated the value of multiple methods through expert assessment of keywords in context, combining human judgment with computational linguistics. Compared to other machine learning approaches, the dictionary one is rather easy to comprehend and use: the dictionary consists of text entries and can be read by anybody just like a traditional dictionary. The easy-to-use software package LIWC (Tausczik & Pennebaker, 2010) or the R package quanteda (Benoit et al., 2018) are available to researchers and can be taken into use with little preparation or cost compared to approaches requiring large computing power and where it is often not clear what happens behind the curtain. LIWC enables users to upload customized dictionaries and select specific text files or directories containing such files – for example, meeting transcripts. The software output indicates the proportion of words from the uploaded dictionary that can be attributed to various dimensions. This efficient process facilitates a comparative analysis of texts based on the theoretically-grounded Multi-Focus Model (Hofstede et al., 1990) with minimal effort.

4) Validation of the Multi-Focus Model on Organizational Culture: We successfully derived the six dimensions of organizational culture based on Hofstede et al. (1990) from various textual sources: employee reviews, mission and vision statements, and annual reports. The dimensional structure resembled the survey-based model in all text corpora, with congruence coefficients between .913 and .976, indicating strong alignment. While these results do not imply the model’s universal applicability, they demonstrate that the organizational culture dimensions, as theorized by Hofstede, are not confined solely to survey research methods. This finding is important as it provides evidence that the Multi-Focus Model is not a methodological artefact of self-report surveys but suitable to describe the (non-)functioning of an
organization more generically, opening the doors to new applications of the model. For example, interested buyers or prospective employees of companies could use the model to gain insights into an organization’s working culture without contacting the company.

**Limitations**

As our study is the first one developing a theory-based dictionary to replicate a survey-based model from various large text data sources, we were only able to cover some aspects, leaving ample space for future research:

It is still uncertain to what extent dictionary-based text analysis results are equivalent to survey results. We quantified the construct of organizational culture from textual data, and we established structural equivalence (similarity of relationships) with the results of the corresponding survey instrument. However, like in similar studies in psychology (e.g., Arnulf & Larsen, 2020; Kjell et al., 2019), we cannot claim metric or scalar equivalence (He & van de Vijver, 2012). Finding the same dimensional structure in survey and text data does not mean that a particular organization would score similarly regardless of whether survey or text data was analysed. Grunenberg et al. (2024) pioneered predicting job applicants’ Big Five scores by analysing CVs and short text responses. Although their tool outperformed recruiter judgements, the average correlation between predicted and self-reported personality traits was only moderate ($r = 25$). In organizational behaviour, predictive accuracy remains unestablished due to a lack of studies with matched text and survey data. Currently, there is insufficient evidence to claim the superiority of either survey or text data analysis.

We found that the dictionary works best on the types of text that were used for its development, like mission and vision statements and employee reviews. Further, limited resources forced us to map only a selection of potential candidate words following the KWIC approach, which has affected the association of the n-grams in the dictionary and the overall dimensionality of the final solution. The DOCP does not claim to include all words potentially related to organizational culture, and we do not know how well the dictionary performs with other types of data that were not used in the development of the dictionary, such as internal communication (emails, chat messages, meeting transcripts). Hence, caution should be exercised when relying solely on one text source for understanding an organization’s culture, as the quality and type of text data are crucial for the quality of the result. Our decision to use several types of data in the development phase addressed this limitation already to some degree.

Our DOCP consists of English words and phrases and hence can only be applied to English texts. While dictionary translation is cumbersome and may not take national culture differences into account, developing a dictionary for every language and national culture is not practical. The translation of source texts into English may be the best approach in most cases under the bag-of-words assumption that the word order is irrelevant and documents are treated as containers (Banks et al., 2018).

Future research could focus on the predictiveness of text analysis results on various indicators such as profitability or sick leaves. The challenge of such a study is not only the gathering of high-quality textual data and the corresponding indicators from the same organizations but also establishing the correct linkage of text data, typically available for certain organizational units within one country, and indicators, often only available on a global level in the case of international organizations.

**Managerial implications**

The findings of this study have several managerial implications for organizations, highlighting the value of text analysis in assessing organizational culture. Firstly, the study demonstrates that organizational culture dimensions can be derived from multiple types of text data, such as employee reviews and mission and vision statements, in an effective and, arguably, more nuanced manner than with employee surveys. This allows managers to gain valuable insights into their organization by analysing these textual sources. Secondly, the research underscores the effectiveness of even relatively simple quantitative text analysis techniques for reproducing models established through quantitative survey methods. This implies that text analysis tools can complement or validate survey-based assessments of organizational culture. By incorporating these text analysis techniques into their assessment of organizational culture, managers can develop a more comprehensive understanding of the cultural dynamics within their organization. Text analysis will also provide a cost-effective and efficient way to continuously monitor and assess cultural trends within the organization, as well as to track the effectiveness of cultural change initiatives over time (Rathje et al., 2023). Managers might utilize this model as a framework to understand and describe the culture of their organization, regardless of the data sources used.

The study highlights the need to explore further the suitability of different text data types for assessing organizational culture. Not all texts are the same. The lower congruence between the results obtained from the DOCP dictionary applied to annual reports (coefficient of .913) suggests that caution should be exercised when relying solely on this text source for understanding an organization’s culture, as communication by management through these outlets apparently does not convey the organizational culture sufficiently. When comparing the DOCP results of employee reviews with the DOCP results of mission and vision statements, different perspectives are compared: the actual organizational culture perceived by the employees versus the projected organizational culture as the organization presents it. Managers should be aware that i) different sources of text provide a different picture of their organizational culture, ii) hence cannot be directly compared between each other, and iii) for a comprehensive insight into the culture, a combination of survey and the analysis of different text sources seems to provide the most accurate description. Analysing different text sources may help identify gaps between the projected and actual culture of an
organizational culture. Managers are best advised to use a mix of survey and text analysis to gain a comprehensive understanding of their organization’s culture.

**Conclusion and perspective**

By adopting a theory-driven approach with multiple sources of text data for reproducing the Multi-Focus Model on organizational culture, this study has contributed to the still scarce literature on the measurement of organizational culture using text. The developed dictionary provides researchers and managers with an exciting new tool for describing organizational culture, providing more comprehensive insights compared to self-report surveys only. Given the similar structure of text and survey results, a takeaway for survey research is that semantic themes largely predetermine associations between constructs such as cultural dimensions (Arnulf & Larsen, 2020; Kjell et al., 2019) – they help us understand how people respond to questionnaires. The measures derived from theory-driven text analysis were found to be at least equally insightful as data-driven text analysis approaches. The LIWC technology (=counting words) is somewhat restrictive for such a complex task as picking up on context in text, and more advanced text analysis tools, especially large language models such as ChatGPT, may soon provide better possibilities for analysing organizational culture from text as recent studies suggest (Fischer et al., 2023; Rathje et al., 2023). However, generic AI tools are not able to analyse text based on established and validated organizational culture theory; they need to be trained. Here, our study and its components, such as the allocated keywords in context, can serve as a basis for the development of AI-based tools. Further, dictionary-based methods will remain helpful for understanding the black box of large language models.

**Notes**

1. Not to be confused with Hofstede’s framework on national culture, the “6D Model” (Hofstede et al., 2010)
2. The open questions of Great Place to Work Germany were: (1) Is there anything unique or unusual about this organization that makes it a great place to work? Please give specific examples. (2) If you could change one thing about this organization to make it a better place to work, what would it be? Please give specific examples. (3) What values does your employer stand for? Please describe in three words what is unique about your employer from your point of view.

**Acknowledgements**

We would like to thank MBA-Exchange.com for the provision of mission and vision statements and Great Place to Work Deutschland GmbH for the provision of employee review data. We would like to acknowledge and express our deep gratitude to Peter Holz, who passed away during the process of writing this paper. He will be greatly missed.

**Disclosure statement**

One of the authors is employed by Hofstede Insights Ltd.

**Funding**

This study has been funded by Business Finland.

**ORCID**

Michael Schachner https://orcid.org/0000-0002-0362-6506
M. Murat Ardag https://orcid.org/0000-0001-9111-1088
Peter Holtz https://orcid.org/0000-0001-7539-6992
Hester van Herk https://orcid.org/0000-0002-3503-5698
Michael Bender https://orcid.org/0000-0003-0300-5555
Klaus Boehnke https://orcid.org/0000-0002-5435-4037
Henrik Dobewall https://orcid.org/0000-0003-3515-2138

**References**


