Modeling Organizational Culture with Workplace Experiences Shared on Glassdoor

Vedant Das Swain*, Koustuv Saha*, Manikanta D. Reddy, Hemang Rajvanshy, Gregory D. Abowd, Munmun De Choudhury
Georgia Institute of Technology
Atlanta, Georgia, USA
{vedantswain, koustuv.saha, mani, hemangr, abowd, munmund}@gatech.edu

ABSTRACT
Organizational culture (OC) encompasses the underlying beliefs, values, and practices that are unique to an organization. However, OC is inherently subjective and a coarse construct, and therefore challenging to quantify. Alternatively, self-initiated workplace reviews on online platforms like Glassdoor provide the opportunity to leverage the richness of language to understand OC. In as much, first, we use multiple job descriptors to operationalize OC as a word vector representation. We validate this construct with language used in 650k different Glassdoor reviews. Next, we propose a methodology to apply our construct on Glassdoor reviews to quantify the OC of employees by sector. We validate our measure of OC on a dataset of 341 employees by providing empirical evidence that it helps explain job performance. We discuss the implications of our work in guiding tailored interventions and designing tools for improving employee functioning.

Author Keywords
organizational culture; social media; glassdoor; wordvector

CCS Concepts
•**Human-centered computing** → *Empirical studies in collaborative and social computing; Social media;* •**Applied computing** → *Psychology;*

INTRODUCTION

> How does your company's leadership measure success? Sales? ROI? That's pretty typical. I don't want to pick on Uber, but its issues should have leadership everywhere asking another question: "How healthy is our culture?" — Taro Fukuyama[1]

Culture is an ether that binds human civilization through its evolution. Culture encapsulates a society's practices, beliefs, attitudes, values, perceptions, rituals, art, philosophy, and even technology [23]. In an organizational context, certain norms and principles that are believed to optimize the workforce and maximize efficiency are referred to as *organizational culture* (OC) [11, 90]. This embodies a core value system which affects the development and execution of new ideas, and the management of unexpected events like crises [25, 89]. While metrics such as revenue and profit are standard methods to gauge the effectiveness of an organization, the culture of an organization is both an indicator and a factor to influence its effectiveness [122]. From the employee's perspective, comprehending OC can help foretell their loyalty and commitment [89] because community can affect human behavior [22, 33].

Organizational studies have employed a variety of survey instruments to quantify OC, but these come with their own challenges [29, 31, 50, 62, 97]. These instruments are limited in scalability and temporal granularity. Besides, conducting such studies in organizational settings leads to unique problems because of employee anxieties regarding the confidentiality of their opinions [7, 81]. Therefore, the workplace context can invite multiple biases, such as response (or non-response) bias, study demand characteristics, and social desirability bias [13].

In contrast, workplace review platforms contain self-initiated and anonymous reports [128] that stand to mitigate many of the biases introduced by survey studies [57]. Glassdoor is one such platform with publicly posted reviews of workplace experiences. Not only do these reviews contain objective information like pay, hours and benefits but also the free-form text that encapsulates various nuances of OC [11, 58]. Take for instance a review that states, *[Company] work was horrible, and upper management is poor at recognizing achievement, but the opportunity to work with my colleagues kept me coming in daily*. The language in this shared experience reflects an organizational culture where recognition is not prioritized but concern for others and cooperation is upheld. In fact, through the affordance of descriptive text, platforms like Glassdoor provide an accessible, scalable and flexible medium to express cultural and ecological differences [51]. Our work leverages the language used in publicly visible employee reviews to computationally model OC and augment our understanding of it. Specifically, this paper has the following research aims:

Aim 1. *To operationalize OC as a multi-dimensional construct and validate it with language on Glassdoor.*

Aim 2. *To computationally model OC of an organizational sector, and evaluate if it explains employee job performance.*

***These authors contributed equally to this work.**

[1] *https://www.recode.net/2017/4/18/15327848/workplace-culture-measurement-metrics-healthy-uber*

CHI'20, CHI Conference on Human Factors in Computing Systems (CHI '20), April 25–30, 2020, Honolulu, HI, USA

DOI: https://doi.org/10.1145/3313831.3376793

Our first research aim strives to build a usable construct of OC, based on Glassdoor data, that captures various aspects like interpersonal relationships, work values, and structural job characteristics. Towards this, we use established frameworks from the domain of organizational psychology [29, 31, 50, 62, 97] to identify job descriptors related to OC and represent them as word-vectors. We ground our approach in the literature, to model organizational culture in the lexico-semantic space of word embeddings [94], and validate this word embedding based construct of OC. This produces a codebook of lexical phrases that closely align with different dimensions of OC.

Next, given a reliable representation of OC we seek to examine if it explains individual performance [89, 90]. We apply our OC construct on Glassdoor reviews and quantify the OC of companies by sector (e.g., management, production, or computer). On a ground truth dataset from the Tesserae project [77, 101] with 341 employees from three companies we find that incorporating a measure of OC improves on intrinsic traits (such as demographics and personality) to explain an employee's task performance and citizenship behavior. This renders empirical evidence that OC explains human functioning and exhibits an application of our construct. Finally, we discuss the implications of this measure of OC for employees and organizations.

Privacy and Ethics. This work is committed to secure the privacy of the Glassdoor reviewers, the company names, and the individuals whose (groundtruth) survey data on individual difference attributes were used. These individuals signed informed consent to provide the survey responses as a part of the Tesserae study, which was approved by the relevant Institutional Review Boards at researcher institutions. Also, despite working with public and anonymized data from Glassdoor, this paper reports paraphrased excerpts of the posts to balance the sensitivity of privacy, traceability, and identifiability, as well as provide the context in readership.

BACKGROUND AND RELATED WORK

Organizational Culture

While "culture" has been interpreted in several ways through unique perspectives across multiple disciplines, *organizational culture* (OC) specifically refers to a socio-cognitive model of emergent standards and norms that help individuals to make sense of their surroundings [16, 113]. Entities, like upper management, often propagate a set of expectations to guide employees in novel and familiar situations [89]. This includes certain assumptions regarding daily interactions at the workplace [113]. Thus, organizational culture emerges from the interplay of top-down expectations and bottom-up norms [29].

Organizational culture can influence innovation, implementation, cooperation, and conflict management within an organization [89]. Apart from organizational outcomes, it can also affect individual functioning. For example, employees in an organizational culture that values them tend to outperform employees in cultures where they feel replaceable [25]. If an organization is founded on toxic or unethical attitudes, it can impact employee morale [122] and in turn, contribute to low employee performance, low retention rate, and low job attractiveness [61]. O'Reilly further states, "culture is critical in developing and maintaining the intensity and dedication among employees" [89]. When employees are empowered and trained to augment their abilities, they exhibit greater job satisfaction [141]. Similarly, organizations that encourage effort and include formal reward structures have fewer instances of workplace misconduct [133]. Research postulates that organizational culture can explain employees' performance and productivity over and above intrinsic traits and abilities [11].

Prior work has used many frameworks to operationalize organizational culture. Some measure it in terms of factors like innovation, competitiveness, decisiveness, and growth-orientation [90]. Others describe it with bidirectional scales like power distance (large/small), uncertainty avoidance (strong/weak), individualism vs. collectivism, or masculinity vs. femininity [62]. Typically these frameworks use survey instruments to measure organizational culture. However, survey measurements are not holistic because of situations where employees are not comfortable to share their opinion, or prefer to have no opinion regarding the organization [7, 81]. Moreover, even when surveys are administered, since it is often within the purview of the organizations, they can be subject to social-desirability biases [13]. Even well-conducted surveys can at times be biased due to lack of candor in responses [13]. Overall, conventional assessments of organizational culture have been reported to lack nuance, context, and applicability in diverse organizational settings [96].

Aim 1. Our approach of operationalizing OC using language circumvents the limitations of conventional evaluations as it harnesses the potential of crowd-contributed, anonymized, and publicly available Glassdoor data. This is grounded in four OC instruments — 1) *Organization Cultural Inventory* (assesses OC with 12 task and interpersonal styles [30]), 2) *Organization Culture Profile* (assesses OC with 54 value statements [90]), 3) *Hofstede's Organization Culture Questionnaire* (assesses OC on 6 independent dimensions) [62], and 4) *Organization Culture Survey* (assesses OC on 6 components [50]).

Aim 2. We draw motivation from Chamberlain's report that employee perceptions on workplace culture (from Glassdoor) share a positive association with a company's financial performance [25]. We examine how quantifying OC helps to explain individual job performance. We measure this on two metrics — 1) *IRB scale* [136] measures In-Role Behavior that characterizes the proficiency at performing appointed activities and tasks, and 2) *OCB scale* [47] measures Organizational Citizenship Behavior which characterizes participation in extra-role activities that are not typically rewarded by the management [17, 82, 91, 109]).

Language and Perspectives on Culture

Language provides a medium to consciously verbalize beliefs and values [11]. Language can reflect differences of personal and situational traits [51]. Therefore, investigating language can identify manifestations of organizational culture [58]. For example, the variations in word use in work emails have been noted to be distinct for individuals who have internalized the culture of a workplace [41]. Similarly, linguistic cues found in internal communications of organizations explain cultural integration between employees [52, 119]. While these studies investigate the acculturation process, they are limited in explaining what culture actually is, or how it affects other outcomes. Moreover, since these approaches harness language

Pros	Cons
1) Great teams 2) Talented co-workers 3) Not stressful 4) Good work-life balance	Most departments offer no flexibility in work schedule. My manager doesn't allow me breaks for doctor appointments, child's school activities
Good work environment, nice people. Lots of fun working on cool technology. Location is also superb.	No communication from upper management, Pay is not nearly as competitive as market salaries.
Friendly, outgoing coworkers. Very health-conscious environment. Activities are encouraged and supported.	Little recognition for overtime hours, no WFH alternatives even with bad weather, poor work-life balance

Table 1. Example paraphrased excerpts in *Pros* and *Cons*.

used in within-organization communication channels (e.g., work email, internal social media), they are likely to dampen candid perspectives of the workplace experiences [13].

Advances in the computational social science has shown the potential of social media and other publicly accessible online data to characterize organizational culture. For instance, textual analysis of annual reports has been used to classify the construct and in turn, explain a firm's risk-taking behavior [85]. Recent work has used Glassdoor to infer individual aspects of organizational culture, such as "goal-setting" [80], or "risk-taking behavior" [85]. Notably, Pasch characterized six dimensions of organizational culture through language, and found a relationship between perceptions of culture and performance of an organization [93]. Similarly, a bag-of-words analysis of reviews for corporate values has been associated with organizational effectiveness [75]. While research reveals the potential of Glassdoor reviews to quantify organizational culture, these works inspect a limited set of dimensions and the implications are primarily organization-centric. In contrast, our work operationalizes this measure based on several dimensions collated in a domain-driven way, and examines its influence on employee-centric outcomes like job performance.

Online Technologies and Workplace Experiences

Online platforms are becoming a powerful resource to study employee activities and interpret their behaviors – a line of research that is extensive in the CSCW and HCI fields [6, 35, 76, 110, 111, 114, 139]. Microblogging at work has been used to build a "common ground" among employees where they can interact with each other's opinions [143]. A case study at IBM found that employees not only share content for a sense of collective identity, but also predominantly express work practices through this stream [126]. Prabhakaran et al. studied power relations in email interactions of employees [95]. Employee engagement on such mediums motivate the exploration of the free-form text they harbor as these the descriptions often explain work based rituals, ideas and beliefs.

Employees use social media for a variety of purposes such as information seeking, knowledge discovery and management, expert finding, internal and external networking, and potential collaborations [6, 142]. Moreover, many organizations even have internal social media platforms [38, 48]. Shami et al. proposed a tool, *Employee Social Pulse* — that analyzed streams of internal and external social media data to understand opinions and sentiment of employees [111]. Similarly, dictionary-based linguistic analysis of such streams has been successfully employed to gauge employee engagement [53, 110]. On the interpersonal side, Muller et al. found that social influence from peers can help shape an employee's engagement [84], and engaged peers seem to be more helpful, enthusiastic, or exude a contagious positive affect [79]. Analyzing employees' social behavior on social media, content endorsement through "likes" is found to represent certain facets of organizational culture [59]. Although such social media has benefits, scholars argue that candid social media use within organizations can be affected by privacy-related concerns such as the breach of boundary regulations and employer surveillance [49, 64, 114].

On the contrary, anonymized platforms like Glassdoor provide "safe spaces" for employees to share and assess their workplace experience [18, 69]. Glassdoor data was used to model brand personalities based on *employee imagery* factors such as working conditions, company culture, and management style [140]. Lee and Kang used Glassdoor data to study job satisfaction and found "Culture and Values" has one of the highest influences in employee retention [72]. Besides, we note that anonymity eliminates the desire *to appear competent and likable*, thereby minimizing deceptive tendencies and enhancing candid and self-motivated / self-initiated posting behavior [45]. Motivated by these findings, we study organizational culture by leveraging a large-scale data source, Glassdoor, which is public-facing and anonymous, but well-moderated.

GLASSDOOR AS EMPLOYEE EXPERIENCE PLATFORM

For our study, crowd-contributed workplace experiences from Glassdoor serve to validate the operationalized OC (Aim 1), and to quantify the OC in an employee sector (Aim 2).

Sharing Employee Experiences

Glassdoor is an online platform (launched in 2008), for current and former employees to write reviews about their workplace experience. As of 2018, there are 57M individual accounts on this platform, and there are 35M reviews posted for 770K companies [120]. Glassdoor reviews require ratings and free-form text. Employees can rate their overall experience on a scale of 1 to 5, and optionally add ratings for fields like career opportunities, compensation, and senior management. The free-form text field requires employees to submit descriptons of their workplace experience, in separate sections for *Pros* and *Cons* (Table 1). This text describes many salient workplace themes, such as work-life balance, management, pay, benefits, growth opportunities, facilities, and interpersonal relationships.

Quality of the Content

In Glassdoor's published community guidelines and norms for content submission, they state that they *strive to be the most trusted and transparent place for today's candidate to search for jobs and research companies* [57]. Both contributing content and consuming content necessitates an individual login. It only allows individual accounts with *permanent, active email address, or a valid social networking account* to submit content, with a maximum allowance of *one review, per employee, per year, per review type* [99]. Glassdoor moderation involves proprietary content-analysis technology as well as human moderators. Any reviews deemed to be incentivized or coerced, are either not allowed or removed from the platform. In addition, Glassdoor offers the option to flag content, which is evaluated on a case-by-case basis. To ensure a non-polarized distribution of reviews, Glassdoor implements a key incentive policy known as, "give to get" [128]. In this model to get full access

Category	Organizational Culture Dimensions
Interests	Conventional, Enterprising, Social
Work Values	Relationships, Support, Achievement, Independence, Recognition, Working Conditions
Wk. Activities	Assisting & Caring for Others, Establishing & Maintaining Relationships, Guiding & Motivating Subordinates, Monitoring & Controlling Resources, Training & Teaching Others, Coaching & Developing Others, Developing & Building Teams, Resolving Conflicts & Negotiating
Social Skills	Instructing, Service Orientation
Struct. Job Characteristics*	Consequence of Error, Importance of Being Exact, Level of Competition, Work Schedules, Frequency of Decision Making, Freedom to Make Decisions, Structured versus Unstructured Work
Work Styles	Concern for Others, Leadership, Social Orientation, Independence, Integrity, Stress Tolerance, Self Control, Adaptability, Cooperation, Initiative, Achievement
Interpersonal Relationships*	Frequency of Conflict Situations, Face-to-Face Discussions, Responsibility for Outcomes & Results, Work w. Group or Team

Table 2. 41 Org. descriptors from O*Net to represent the dimensions of OC. The category column indicates the O*Net category of the descriptors. Categories with '*' are subcategories within the "Work Context" cateogry. The table in supplementary material provides a detailed description of job descriptors with the validation source.

to all reviews, viewers must contribute their own review. This paradigm encourages more neutral opinions to be recorded and diminishes the biases of self-selected users [26]. The content posted on Glassdoor remains anonymous, and the moderation strategies ensure that no sort of individual-identifiable detail is disclosed in the content. However, each review comes tagged with the reviewer's role, employment status (current or former), and location of employment.

AIM 1: OPERATIONALIZING ORGANIZATIONAL CULTURE

In order to measure OC through language on Glassdoor reviews, we need to first operationalize it based on language. We adopt a three-step approach to achieve this: 1) Identify descriptions of multiple dimensions of OC. 2) Transform the descriptions into word-vectors to capture their linguistic and semantic context, so as to represent OC as a collection of these vectors. and 3) Compare the word-vector based OC construct to filter Glassdoor posts related to OC and qualitatively investigate the posts' keywords to establish face-validity.

Identifying Descriptors of Organizational Culture

Language used by a community (or organization) provides a unique lens to interpret its culture [11, 51]. To understand the extent to which a text expresses OC, we first need an established ontology of job aspects that are indicative of different OC dimensions. For this, we obtain job aspect descriptors from the Occupational Information Network (O*Net). O*Net (*onetonline.org*) is an online database of occupational information developed under the sponsorship of the U.S. Department of Labor/Employment and Training Administration (USDOL/ETA). These descriptors are motivated by organizational research literature [54, 60, 123], and are regularly updated with changes in socio-economical and workforce dynamics. O*Net describes 189 different job descriptors, categorized in 17 sub-categories, which are further grouped into 8 primary categories. Each of the 189 job descriptors, like *Stress Tolerance*, *Level of Competition* and *Independence*, is accompanied by a description.

However, not all descriptors necessarily explain OC. For example, descriptors like *Staffing Organizational Units* and *Pace Determined by Speed of Equipment* simply describe characteristics of the job role, not the underlying concept of OC. Therefore we first verify which descriptors align with established frameworks of OC that are widely used in organization research. Two coauthors familiar with organizational studies independently inspected each of the 189 descriptors in O*Net on the basis of four OC instruments, *Organization Cultural Inventory* [30]), *Organization Culture Profile* [90]), *Hofstede's Organization Culture Questionnaire* [62], and *Organization Culture Survey* [50]). Any discrepancies ($n = 23$) with respect to the validity of a job descriptor was resolved by both authors on agreeable themes and concepts. Overall this procedure had a Cohen's κ (inter-rater reliability) score of 0.89 This process retains 41 descriptors, each of which describes an aspect of OC (see Table 2). Also note that these dimensions are not necessarily mutually exclusive or disjoint [90, 100], and we expect a significant overlap in our ensuing analysis. Our domain-driven approach validates the O*Net descriptors on the basis of multiple different frameworks because no single conceptual framework describes OC exhaustively [29, 100].

Transforming Descriptors into an OC Construct

While O*Net provides explanations of the 41 descriptors of OC, simply tokenizing the keywords in these descriptions would not adequately capture the concept of OC. Therefore to address this challenge, we encapsulate the linguistic and semantic context of these descriptions by using the concept of word embeddings [44, 105]. This approach represents words as a vector in a higher dimensional space, where contextually similar words tend to have vectors that are closer. In particular, we use pre-trained word embeddings in 50-dimensions (GloVe: trained on word–word co-occurrences in a Wikipedia corpus of 6B tokens [94]). Building on prior work of representing job aspects in lexico-semantic dimensions [103], we transform the explanations for each of the 41 descriptors (Table 2) into a 50-dimensional word-embedding vector. These 41 word-embedding vectors essentially characterize multiple dimensions of OC in a latent semantic space. Collectively, they constitute our operationalized construct of OC.

Validating our Operationalization of OC

While our operationalization of OC captures the information contained in 41 descriptors (obtained from O*Net and validated from domain assessments of OC), we need to establish its validity for practical use. We qualitatively inspect the top keywords in text from our Glassdoor dataset that is relevant to OC. We elaborate on this procedure in the following segments.

Compiling the Glassdoor Dataset

To obtain a diverse but voluminous dataset on Glassdoor, we first consult the *Fortune 500* list (ranked by revenue) [1] and obtain the top 100 ranked companies. Since only 8 of these companies appear in the list of *Fortune 100 Best Companies to Work For* [28] we believe our sample is not dominated by companies with positively-skewed employee experiences.

We obtain the public reviews of these organizations using web scraping. For each review, we collect the textual components (segregated into *Pros* and *Cons*) and the reviewer's employment information (role and location). Table 1 shows three

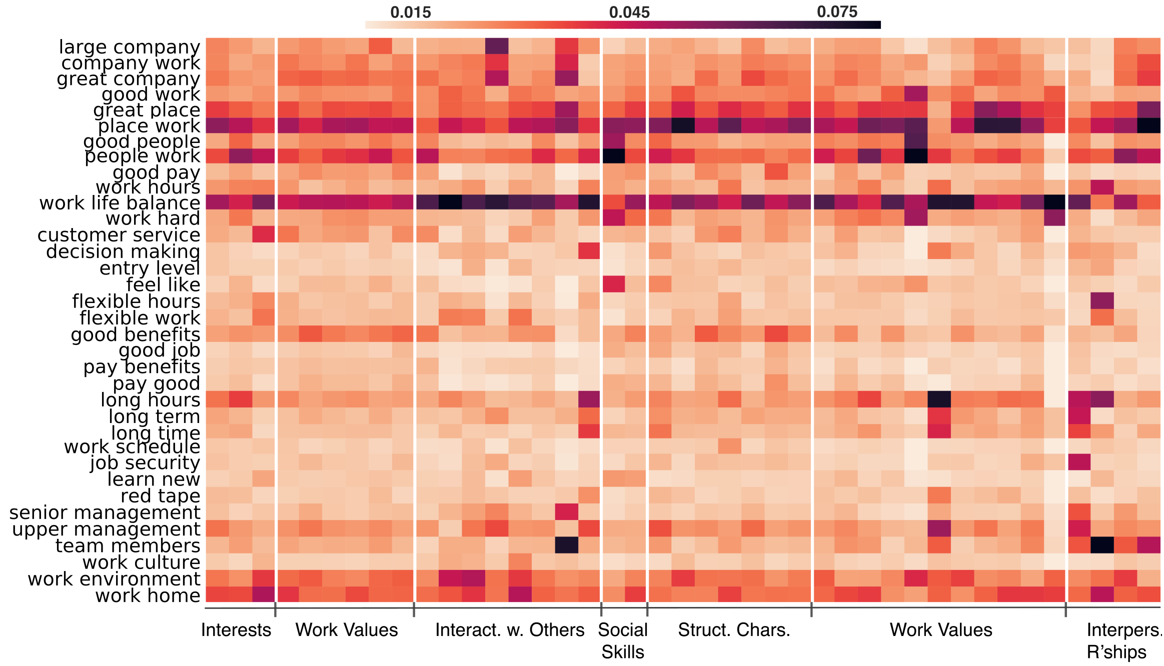

Figure 2. Top n-grams in sentences about OC (excluding lexical variants of keywords). Darker colors (higher TF-IDF score) indicate greater relative importance within a particular dimension. Dimensions have been categorized corresponding to the scheme in Table 2

Measure	Total	Mean	Stdev.
Reviews	616,605	6,702	8312
Pros Sntncs.	1,386,787	15,073.77	18,408.64
Pros Words	10,747,265	17.42	20.91
Cons Sntncs.	1,715,875	18,650.82	22,786.10
Cons Words	17,150,342	27.81	47.24

Table 3. Descriptive stats. of Glassdoor dataset of 92 companies (sourced from top 100 of Fortune 500). Aggregated values are per company.

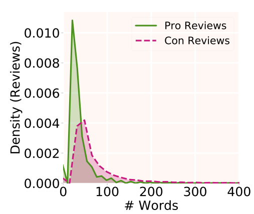

Figure 1. Dist. of no. of words per review in the Glassdoor dataset of Fortune 100 companies.

Example Text	OC Dimension
Great training, really genuine and supportive colleagues, great ways to get involved with interest groups—Proposal writing, research for new industry areas, volunteer activities	Social
In many instances rank was invoked just to prove a point, rather that using data for the same.	Importance of Being Exact
The drive to succeed is key, however, it's not a cut throat competition - people are humble and people at all levels are interested and willing to develop those at the lower career levels.	Level of Competition
If you have a goal and willing to work on it, senior management will have a genuine interest in helping you succeed.	Coaching and Developing Others
A lot of emphasis is on firm activities making it difficult to build relationships as you can only meet coworkers on Fridays, if they do come.	Establishing and Maintaining Interpersonal Relationships
New recruits are immediately given responsibility, and can take complete charge of their career development.	Initiative
Lot of group work makes the work easier and more fun.	Independence

Table 4. The word-vector representation of these sentences that show a cosine similarity of 0.90 or greater for the corresponding OC dimension. Note that the same sentence can reflect multiple dimensions, but we only list one for brevity.

example excerpts in *Pros* and *Cons* components. In sum, we obtain 616,605 reviews from 92 companies (at the time of writing 8 companies did not have profiles on the platform) that were posted on Glassdoor between February 20, 2008 and March 22, 2019, amounting to 10,747,265 words in the *Pros* segment and 17,150,342 words in the *Cons* segment (ref: Table 3 and Figure 1). We note that the content distribution is skewed towards the *Cons*, but this observation aligns with activity on other review platforms [63]. Despite the possibility that some of these reviews could be capricious and circumstantial, this work intends to leverage the ample volume of data and capture themes at an aggregated level. Additionally, all our computation normalizes data by volume.

Filtering Posts about Organizational Culture

First, we derive a word-vector representation of every sentence in the 616,605 posts (∼3M) from the Glassdoor dataset. Next, we use cosine similarity to measure the similarity between each sentence's word-embedding representation and each of the 41 dimensions of OC [10, 104]. Higher cosine similarity indicates that the sentence is semantically similar, or "talks about" that particular dimension of OC. We retain any sentence that exhibits a similarity of more than 0.90 with any of the OC dimensions. Note that the same sentence may express an opinion about multiple classes; for example, a post reading "Some staff is able to negotiate to avail work from home at least one day per week" relates to *Work Styles: Social Orientation*, *Work Values: Relationships*, and *Work Values: Independence*. Table 4 enlists a few paraphrased examples.

Establishing Face and Construct Validity

Since the sentences that clear the threshold only relate to OC through the latent semantic space of word-embeddings we now want to investigate the actual language used in the content. We obtain the top 100 keywords (n-grams, n=2,3,4) in all sentences (above the similarity threshold of 0.90). Then, we compute the TF-IDF score for these keywords across each of the 41 OC dimensions (similar to [105]). Essentially, this helps us uncover the importance of each keyword in the sentences that refer to an aspect of OC. Figure 2 visualizes the relative importance of these keywords (the supplementary document provides a heatmap with all top 100 *n*-grams). We draw upon the validity theory [87], to establish face and construct validity of contextualizing OC in Glassdoor data by qualitatively examining the importance of the keywords in the OC dimensions.

The most dominant keyword across several dimensions is **work life balance**, and its lexical variants like "life balance", "work life". This recurrence could be because notions of work–life balance has many facets (beyond work-family conflict) such as personal needs, social needs and team work [92]. For instance, this *n*-gram is important to the *Social* dimension of OC because it characterizes altruistic behaviors and aid of colleagues [61]. Similarly, dimensions like *Assisting and Caring for Others*, *Coaching and Developing Others*, and *Training and Teaching others*, inherently overlap with the team based aspect of "work life" [50, 97]. Socially supportive and inclusive workplaces tend to foster better work–life balance, these key-words co-occur with language referencing social and interpersonal dimensions, for example *"[Company] tries to ensure work life balance, whether it works is another story as everyone seems too dedicated."* and *"[Company] offers the best work life balance and true diversity among big firms"*.

Certain keywords are relatively more discriminatory between OC dimensions. For instance, the keyword **"good benefits"** is most important in reviews about dimensions like *Support* and *Recognition*. For employees, reward systems within companies garner reciprocal loyalty and increase the perceived organizational support [43]. For example in this post, "There is effective communication from senior management along with a good benefits package, cutting-edge technology, and a culture of integrity and innovation that provides a very satisfying environment.". Another such keyword is **"job security"**, which is most relevant to experiences that refer to the *Frequency of Conflict Situations* dimension. This draws from the fact that employees in workplaces that have high disagreements require more security and stability of employment [62]. Other examples of identifiable *n*-grams are **"flexible hours"** or **"flexible work"**. These keywords gain maximum importance in text associated with the *Face-to-Face Discussions* dimension. Prior research found that teams with fluid hours accommodate more interactions [20]. Similarly, the terms **"long hours"** and **"long time"** are important in texts related to the *Stress Tolerance*. Longer working hours not only causes fatigue but also increases an employee's exposure to work-related stressors [21, 66, 118], such as that expressed in, *"Client projects can require* ***long hours*** *on short notice, and the general environment can be very demanding and not forgiving."*

Apart from those discussed above, some of the *n*-grams correspond to the dimensions of OC more intuitively. For example, **"good people"** is most important in texts associated with *Resolving Conflict* (*Interacting with Others*), **"senior management"** is relevant to texts about *Frequency of Conflict Situations* (*Interpersonal Relationships*), and **"team members"** dominates experiences about the *Face-to-Face Discussion* dimension (*Interpersonal Relationships*). The evidence we have provided indicates that the OC construct built from curated O*Net job aspect descriptors can capture the OC-related language used in Glassdoor reviews.

Measure	Scale	Range	Mean	Stdev.	Distribution
		Independent Variables			
Demographics					
Age		21-64	34.15	9.01	
Gender	Categorical: Male | Female				
Job characteristics					
Tenure	Ordinal: 10 values [<1Y,1Y,..>8Y]				
Supervisory Role	Categorical - IT | Non IT				
Personality Traits (BFI-2)					
Extraversion	1-5	1.67-4.91	3.42	0.68	
Agreeableness	1-5	2.08-4.91	3.85	0.54	
Conscientiousness	1-5	1.92-5.00	3.90	0.65	
Neuroticism	1-5	1.00-4.67	2.44	0.75	
Openness	1-5	1.17-4.91	3.79	0.60	
Executive Function (Shipley)					
Crystallized: Abs.	0-25	0-23	17.11	2.97	
Fluid: Voc.	0-40	0-40	33.06	3.93	
		Dependent Variables			
Job Performance					
IRB	7-49	20-49	44.48	4.57	
OCB	20-100	32-100	56.20	10.28	

Table 5. Summary of individual attributes for Aim 2.

AIM 2: MODELING OC AND EXAMINING ITS RELATIONSHIP WITH JOB PERFORMANCE

Prior work in the domain states that organizational culture (OC) influences individual performance in the workplace [133, 141]. This motivates us to apply our 41-D model OC on posts of an organizational community (such as occupational sector) to explain the job performance of employees belonging to the same group. In this section, we describe a methodology to computationally model OC with our proposed construct. Then, we evaluate whether our proposed model can augment our understanding of employee-functioning beyond what is explained by individual differences.

Compiling the Groundtuth Dataset

Towards our Aim 2, we obtain a groundtruth dataset of individual job performance from three companies C_1, C_2, and C_3, and the Glassdoor reviews of these three companies. Our groundtruth dataset comes from the Tesserae project, a large-scale multi-sensor study that recruited information workers in cognitively demanding fields (e.g., engineers, consultants, managers) [77, 78, 101]. This provides us the individual difference attributes and job performance of 341 information workers across 18 unique sectors in three companies C_1, C_2, and C_3 in the U.S. Table 5 summarizes the distribution of all these measures across the 341 individuals in our groundtruth dataset. The individual attributes include demographic details such as age, gender, education, supervisory role (supervisor / non-supervisor), income, and their role in the organization. This dataset also contains information on personality traits and executive function, both of which are robust indicators of job performance. The Big Five Inventory (BFI-2) scale [117, 125] measures personality traits across the big five personality traits of openness, conscientiousness, extraversion, agreeableness, and neuroticism. The Shipley scale [112] measures the executive function in terms of fluid and crystallized intelligence.

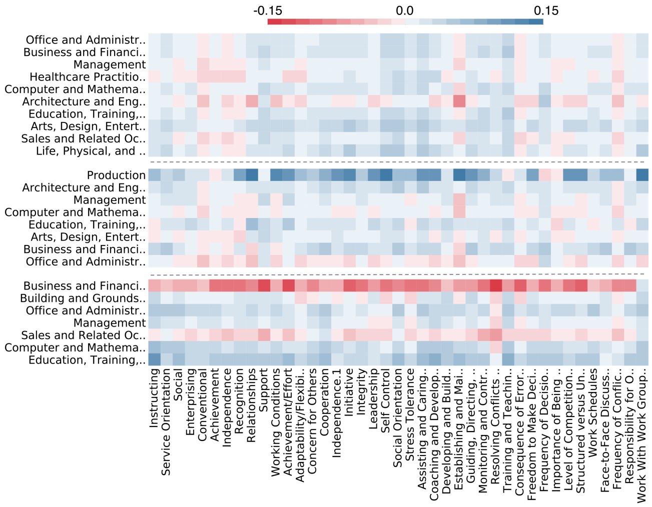

Figure 3. Organizational culture as quantified via Glassdoor data per organizational sector in three companies *C_1* (top), *C_2* (middle), and *C_3* (bottom). The color and intensity of the cells represent the positivity or negativity in that dimension of organizational culture.

The dataset provides two job performance measures *the IRB scale* [136] (In-Role Behavior) and *the OCB scale* [47] (Organizational Citizenship Behavior).

The dataset classifies participants into 18 unique sectors based on role. The top three sectors by participant count are "business and financial operations" (115), "computer and mathematical" (105) and "management" (50), but the dataset also features sectors like "office and administrative support" and "healthcare practitioner". We find 25 combinations of company and sector (eg. {C_1, Computer and Mathematical}, {C_2, Management and Consultancy}, etc.). Corresponding to the same companies (C_1, C_2, and C_3) and the same sectors, we obtain 23,791 reviews on Glassdoor (22,794 for C_1, 574 for C_2, and 423 for C_3). At an average of 350 reviews per {*company, sector*} group. These reviews contain 1,654, 134, and 108 unique roles respectively that mapped to the 18 sectors. For this, we use a semantic similarity based approach using pre-trained word vectors (trained on 6B tokens on the entire Wikipedia corpus) [94], and next, two researchers manually verified the mapping, and edited the sector label wherever necessary.

Modeling and Quantifying OC by Org. Sector

Since culture is a collectively experienced and manifested, we consider experiences expressed by employees who share a common basis, such as a team, department or sector in an organization. Such an approach facilitates a robust and replicable mechanism to study OC both between and within organizations – as prior work investigated the phenomenon on varying levels of organizational granularity [40, 113]. In as much, we are motivated by recent social media language analyses that use word embeddings [103, 104] to model OC.

We first collate all the reviews posted in different company sectors. Then, using word-embedding based cosine similarity, we obtain the similarities of every review sentences with each of the 41 OC dimensions. We cannot simply apply the similarity measure directly as certain posts could be talking about a dimension either positively or negatively. Consider the *Independence* dimension (in *Work Styles*), which refers to a culture that expects employees to be unsupervised and self-motivated. For some employees, such a culture can be favorable, while for others it can become a challenge. Therefore, we qualify the raw similarity score between a post with the help of Glassdoor's *Pros* and *Cons* structure. We assign a weight of $+1$ to those sentences labeled as *Pros* and -1 to those sentences labeled as *Cons*. We obtain the weighted average of cosine similarities for each dimension. Together, this represents a 41-dimensional vector of OC, where a value per dimension is equivalent to how positive or negative that dimension is lexico-semantically spoken about in an organization's Glassdoor reviews. In this way, we can describe the OC of any group of employees in terms of a 41-D vector as long as we can retrieve a corpus of Glassdoor like experiences.

Figure 3 shows the distribution organizational culture in 41 OC dimensions in our dataset. We observe that OC varies across sectors both within and between companies. For example, the reviews from employees in the sector "business and financial operations" shows contrasting trends — while the reviews in C_1 and C_2 talk about OC in a similar way, the reviews of C_3 typically discuss dimensions of OC in *Cons*. We note that company characteristics of the scale and varying interests of employee-base could influence these sort of differences in the employee perspective on culture [37, 113].

Relationship between OC and Job Performance

As human behaviors are affected by the complex interplay between an individual and the culture they are embedded within [22], we hypothesize that our approach of operationalizing OC can explain an individual's job performance which we obtain at the beginning of this section [89, 90, 141].

> **Hypothesis.** Organizational culture provides significant explanatory power towards one's job performance.

We test our hypothesis by predicting job performance — 1) In-Role Behavior (IRB) and 2) Organizational Citizenship Behavior (OCB). We first build a baseline model (*Model 1*), with individual attributes, to predict job performance (*Model 1*). This is motivated by prior work that extensively established that individual difference attributes (such as demographics, personality, and executive function) are strong indicators of job performance [12, 37, 55, 78, 98, 106, 108]. We also control for the individual's organizational sector. Next, we build an experimental model (*Model 2*), where we incorporate OC alongside the individual difference variables, and predict the same job performance measures again (*Model 2*). Here we include the 41-D representation of OC based on the Glassdoor posts of each employee's {*company, sector*}. If *Model 2* is better (statistically significant) in explaining the job performance measures than *Model 1*, then our hypothesis holds true.

$$JP \sim gender + age + income + supervisory\_role + tenure + exec.\_function + personality + org.\_sector \quad (Model\ 1)$$

$$JP \sim gender + age + income + supervisory\_role + tenure + exec.\_function + personality + org.\_sector + OC[41D] \quad (Model\ 2)$$

Since the job performance measures are continuous, both models are regression estimators. We use three types of linear regression models with regularization, Lasso (L1 regularization), Ridge (L2 regularization), and Elastic Net (both L1 and L2 regularization), and two non-linear regression models, SVM regressor and Random Forest regressor. To tune the parameters

Measure	IRB		OCB	
	Model 1	*Model 2*	*Model 1*	*Model 2*
Algorithm	Lasso	Ridge	Ridge	Ridge
R^2	0.23***	0.28***	0.15***	0.24***
Pearson's r	0.43***	0.45***	0.32***	0.41***
SMAPE	3.67	3.65	6.96	6.71

Table 6. Summary statistics of the "best" regression models in *Model 1* and *Model 2*, where *Model 2* includes organizational culture, whereas *Model 1* does not. ***: $p < 0.0001$

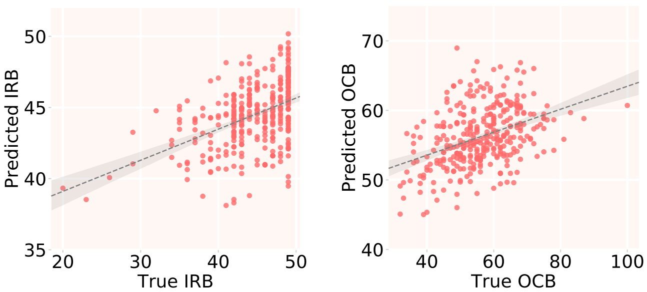

Figure 4. Scatter plots showing true and predicted values per *Model 2* of IRB (left) and OCB (right).

of the models, we use a grid search [116]. We use a leave-one-out (*loo*) [2] methodology to train and predict over our dataset, that is, we iteratively train our models with one held-out data sample, and predict on that held-out sample. Finally, we collate all the predicted data points, and obtain the pooled model performance measures — these include Pearson's correlation and Symmetric Mean Absolute Percentage Error (SMAPE) to evaluate the predictive accuracy of our models, and R^2 to evaluate the model fit (here JP is job performance).

Does Organizational Culture Explain Job Performance?

Model Performance. Table 6 summarizes the fit and accuracy metrics of *Model 1* and *Model 2* for predicting job performance measures (IRB and OCB) (see Fig. 4 for scatter plots). First, we observe that the data behaves linearly, as neither SVM regressors and Random Forest regressors performs better than the linear models. Next, and more importantly we find that the *Model 2* which include the organizational culture as an independent variable (or feature), performs better than the *Model 1*. In the case of IRB, for instance, *Model 2* fits 22% better, and *Model 2* predicts with 5% better-pooled correlation, and 0.6% lower SMAPE. In the case of OCB, the improvement is significantly high, with 60% better fit, 28% predicted correlation, and 4% lower predicted error compared to the performances on the job proficiency measures given by *Model 1*. All these models fit and predict with statistical significance ($p < 0.01$).

Model Validity. Despite *Model 2* performing better, we need to reject the possibility that this is by chance. We aim to reject the null hypothesis that a randomly generated 41-D vector will perform better than our particular 41-D OC (*Model 2*). Drawing on permutation test approaches [5, 102], we run 10,000 permutations of randomly generated OC vectors. We find that the probability (p-value) of improvement by a randomly generated vector is 0.0002 for IRB and 0.0001 for OCB. This rejects

[2] Our rationale to use *loo* validation over more standard k-fold cross-validation rests on the bias-variance tradeoff [134]. Given the small size of our dataset (n=341), such an approach leads to unbiased but high-variance models per fold. This ensures greater stability, robustness, and reduced randomness in sampling [68, 137].

IRB		OCB	
Variable	**Coeff.**	**Variable**	**Coeff.**
Freq. of Conflict Situations	0.59	Adaptability/Flexibility	-49.92
Service Orientation	6.31	Work Schedules	1.45
Recognition	24.10	Face to Face Discussions	0.36
Independence	-9.93	Importance of Being Exact	-0.46
Responsibility for outcomes	0.89	Coaching Others	-37.43
Working Conditions	-8.58	Instructing	-36.56
Freq. Decision Making	-10.80	Wk. w/ Work Group	-0.003
Enterprising	0.96	Conventional	-167.92
Monitoring Resources	0.80	Support	-72.41
Initiative	-9.20	Maintain Relationships	75.35

Table 7. Predicting job performance by *Model 2*: Summary of top 10 regression coefficients ranked on variable importance [56].

the null hypothesis and reveals statistical significance in the observed improvement by including OC based on our quantification. Further, ANOVA tests to compare *Model 1* and *Model 2* reveals that *Model 2* fits significantly better (p<0.001) for both IRB (F=974) and OCB (F=310). Therefore, supporting our hypothesis, we find that OC as computationally modeled using Glassdoor reviews per organizational culture explains job performance of individuals in those organizational sectors.

Interpretation of Results

Table 7 reports coefficients of the top 10 OC dimensions (ranked on variable importance [56]) in *Model 2*. In-Role Behavior (IRB) assesses an employee's efficiency in accomplishing formal task objectives directly pertaining to their appointed job role. The positive relationship between *Recognition* and IRB is obvious because proficiency in one's assigned role leads to rewards through incentive and upward mobility [132]. *Responsibility for Outcomes and Results* is also positively related to IRB because individuals high in conscientiousness are known to be superior in task performance [8, 34, 73]. Experiences talking about *Frequency of Conflict* in the *Pros* more often correspond to higher IRB scores because conflicts(interpersonal, process-based or task-related) are detrimental to performance [65].

Organizational Citizenship Behaviors (OCBs) are not related to formal job roles and typically involve serving the community with extra-role tasks. We find the *Adaptability/Flexibility* dimension to be negatively associated with OCB because an OC which is more open to variable work styles triggers reduced face time between employees leading to fewer opportunities to give back [130]. This also explains why dimensions like *Work Schedule* and *Face to Face Discussions* are positively related to OCB. Another quality of OCBs is that they are based on mutual respect and reciprocality [32, 131]. This explains the positive relationship with experiences favorably describing the *Establishing and Maintaining Interpersonal Relationships* dimension. Additionally, work environments with high job autonomy elicit more OCBs as employees are empowered to use their time for altruistic outcomes [15]. In as much, we observe a negative relationship with the *Conventional* dimension, which represents clear of authority and rigidity.

Post-Hoc: Does Language tell us more than Ratings?

Finally, after we establish that quantifying OC with Glassdoor posts of a sector *does* significantly explain individual performance at workplace, we revisit the question, "is quantifying via language actually effective?" As Glassdoor is a platform that allows individuals to provide ratings, we examine if features based on linguistic aspects of the content offer anything more than raw scores. We build a third model where we only replace OC in Model *Model 2* with mean aggregated rating per sector. The Ridge model performs the best in both the job performance measure predictions. For IRB, this model shows an adjusted R^2=0.24, Pearson's r=0.43, and SMAPE=3.65. For OCB, this model shows Adj. R^2=0.14, Pearson's r=0.32, and SMAPE=6.95 — this model performs only as good as *Model 1* (ref: Fig. 5). So, Glassdoor content when quantified in the lexico-semantic space bears greater explanatory power in comparison to a single numeric rating. This adds credence to our approach of operationalizing o OC as a multi-dimensional construct [90] instead of relying on a single value.

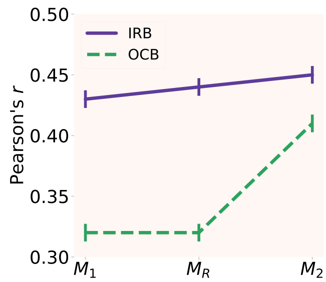

Figure 5. Pearson's r of Models predicting individual job performance (M_1: Model w/o OC, M_R: Model w/ Org. Sector wise Rating, M_2: Model w/ OC via Language)

DISCUSSION

This paper presents a novel methodology to quantitatively model OC through crowd-contributed employee experiences at workplaces. We demonstrate that crowd-contributed workplace experiences on anonymized review platforms such as Glassdoor explains the lexico-semantics of OC. Further, we reinforce concepts in organizational behavior research by validating that this model can significantly explain individual performance at the workplace. We discuss the contributions of our work in understanding workplace experiences, and in designing empirically guided data-driven technologies to help improve organizational functioning.

Theoretical and Methodological Implications

Beyond Surveys/Ratings. By leveraging crowd-contributed experiences of workplaces shared in an unprompted way online, we mitigate the limitations of traditional surveys in assessing OC [29, 31, 50, 62, 97]. While traditional surveys that summarize information into a singular score have its benefits, this compression of information loses the nuance of the multi-dimensional nature of OC [96]. We tackle this challenge by conceptualizing OC on the basis of 41 dimensions and their lexico-semantic space. Another challenge of traditional surveys at the workplace is their vulnerability to a number of response biases [9, 13, 129]. In an organizational setup, a participant's privacy insecurities of getting exposed to management are amplified [49, 64]. This leads to both social desirability and non-response bias. Moreover, many individuals with counter-views and unpopular opinion do not end up participating in such studies to begin with, unevenly skewing the data that is sampled. Alternatively, we use data from Glassdoor where content is public, anonymous, and not actively solicited [57]. Although prior experience and personality can affect public disclosure online, here it is primarily driven by altruism, knowledge, and self-efficacy [27, 70, 71]. Finally, surveys are limited by when and how frequently they are conducted. Although arguably, OC is an enduring construct, it does evolve [4], as incoming new recruits seek to adapt to expectations or bring their own. Moreover, OC can be intentionally changed to improve employee retention and revenue [42, 67, 107]. Unlike surveys, our method can be modified to analyze splices in time and empirically trace the cultural evolution of organizations [121].

Organizational Culture as a Linguistic Construct. Our work contributes a word-vector based lexico-semantic similarity approach to model OC, furthering earlier approaches that use bag of words, topic models, and dimension reduction [2,46,85,93]. While textual data obscures information of momentary physical and social interactions, lexico-semantics of language capture the underlying cultural setup of an organization [11,51,58]. Although review platforms have traditionally been considered as a mechanism to rank, compare, or recommend across entities like companies, our work provides evidence that anonymized (but well-moderated) platforms such as Glassdoor can be leveraged as a reflection of offline and/or situated communities [14]), and their norms and practices.

Practical and Design Implications

Interest in the topic of human resource management is still nascent in the HCI, and cross-disciplinary literature pertaining to workplaces and technology provides several use cases urging the attention of designers [34, 111]. Along these lines, our work presents opportunities to design for personnel management and organizational decision making.

"How is It Like Working in Company X?" Modeling OC can render a normative "signature" of an organization, which can feature on public online platforms. This can help both job-seekers and existing employees to reflect on the assumptions and expectations of a work environment [35,111]. After all, OC has been known to be associated with job attractiveness [19]. Additionally, enterprise-based social-networks, as well as collaborative knowledge bases or "wikis" are already heavily used by new recruits and other geographically distanced employees to learn about and engage in their organization's culture [127]. Since knowledge seeking on such platforms is very high [3,24], integrating language-based models of OC, like we present here, in them can help teams understand the work environment and beliefs and attitudes within an organization.

"How Healthy is Our Culture?" From the employer's perspective, an actionable representation of OC, delivered through privacy-preserving, employee-aware technologies and interfaces, can provide a concrete sense of both individual and collective performance. OC has always been argued to have a strong influence on employee behavior [11, 90]. Yet few companies try to understand their underlying culture, and such efforts are largely limited to gauging a coarse notion of OC [25, 111]. Our approach helps to quantitatively assess OC. Moreover, its multi-dimensional nature can allow a company unpack the atmosphere developing within the workplace through questions like: "Does our culture support work-life balance? Does our organization enhance employee creativity? Or is it concentrated only on productivity? Do we celebrate, incentivize, reward, recognize individual efforts well enough? Do we have enough collaboration? Do employees enjoy doing that?" Importantly, with an ability to quantitatively gauge OC, companies can inspect how well leadership structures model behaviors that embody the company culture, how important events (e.g., IPOs, product releases), may impact the culture, and what steps might address issues of unhealthy culture.

Ethical Implications and Considerations

Meaningfulness of Glassdoor Data: Bias and Abuse. Although our method is agnostic to the nature of the platform, it is undeniable that its credibility and consequences in a practical deployment hinge on the characteristics of the platform. Glassdoor claims to be equitable in its moderation and presentation of different reviews irrespective of ratings [57]. Even though they champion free speech, they avoid platform abuse and illegitimate skewing / polarization of reviews, they establish strict user limitations; *Each individual is allowed one review, per employer, per year, per review type*. However, guidelines can be breached and even updated. And, even with clear guidelines, users can develop behaviors that are "within the rules" but may deter the overall meaningfulness of the data. Moreover, despite content balancing policies like "give to get" [128], review sites like Glassdoor can face *retaliatory utilization* – where dissatisfied employees are more likely to post [63]. A similar problem is intentional tarnishing of employers by trolls. Glassdoor uses an email verification process, but their procedure does not validate an employee's claimed connection to an employer [86]. Since our work shows the applicability of data from platforms like Glassdoor to understand offline organizational constructs, it should motivate stronger policies to avoid misbehavior and presence of "bad actors".

Reputation Building and Divergent Views. In a data aggregation based method like ours, is a small organization as empowered as one with a large amount of users and history? Companies with more employee reviews will be robust to diverse opinions and therefore may find it easy to build and maintain their reputation on public platforms. We also recognize that smaller companies, especially those in early stages, may find it challenging to build a reputation when it can be easily misconstrued with a few extreme reviews. In fact, given the quantitative representation of OC offered by this paper, potential employees could leverage inappropriate portrayals of smaller companies' cultures as an extortion tactic to negotiate pay and benefits [124, 138]. Our work encourages platforms like Glassdoor to consider new ways to protect organizational profiles up until they reach a critical mass of reviews.

Manipulative Intent to Alter Cultural Perception. Formulating the culture of an organization through our approach as is, ignores the nuances of user behavior on sites like Glassdoor [39, 114, 142]. Admittedly, these vulnerabilities can be exploited to harm a company's reputation, and alternatively organizations may game the system to boost attractiveness. Crowd-contributed platforms in other spheres like service and product feedback are rife with problems of "review fraud", where reviewees appoint artificial reviews to alter their public perception [74, 83]. Our model can be abused to selectively manipulate information and jeopardize employee agency and rights, such as by encouraging posts that ignore less desirable cultural attributes, and consequentially harm, or even socially alienate the employees who identify with those attributes.

Limitations and Future Directions

The approach proposed in this paper demonstrates how to operationalize the manifestations of organizational culture in anonymized publicly available posts. Even though the demographics of Glassdoor users are fairly distributed [120, 135], the motivation to generate content is not equivalent for all actors on such platforms [3, 24]. Therefore, it is important to acknowledge that the experiences shared on Glassdoor are only opinions of an underlying organizational culture that may have been influenced by the author's outlook and history. Research in this space needs to consider incorporating this diversity between the authors and their rationale of disclosure.

Next, the primary resource used to construct an objective representation of organization culture is from O*Net which is developed by the U.S. Department of Labor [88]. Additionally, the dimensions of organizational culture captured in our model are motivated by frameworks conceived based on the United States workforce ecosystem. In a cross-cultural analysis, Denison et al. found many of the cultural dimensions associated with high performance in North America did not hold for organizations in Asia [36]. Since an organization is often part of a broader socio-demographical ecology, the geographical culture it is situated in will interact with the culture it fosters within it [115]. Further research with a more diverse organizational sample will help decipher these effects.

Finally, this paper is essentially a feasibility study to establish the utility of descriptive workplace experiences to computationally model organizational culture, and to test if such a model can statistically explain job performance. Even though the volume of employee reviews we analyze is unprecedented, these models are built and evaluated only on a specific set of organizations (e.g., high revenue companies and companies in our ground truth dataset). While this calls into question the generalizability of our results beyond this dataset, our work fosters opportunities to analyze such data for other companies.

CONCLUSION

We empirically studied organizational culture by leveraging large-scale employee-contributed workplace experiences posted on Glassdoor. We examined the linguistic dynamics in public-facing anonymized reviews to describe culture, and developed a theoretically-grounded rendition of it as a codebook. Next, we developed a lexicon to encapsulate culture based on 41 dimensions. We modeled organizational culture for company sectors and tested its explanatory power in predicting employee performance, where we found that our formalization of organizational culture significantly explains individual performance and citizenship behavior, beyond individual intrinsic attributes (eg., demographics and personality). This work bears implications in designing individual- and organization- facing tools to improve organizational functioning.

ACKNOWLEDGMENT

This research is supported in part by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), via IARPA Contract No. 2017-17042800007. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of ODNI, IARPA, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright annotation therein. We thank the Tesserae team for their contributions in realizing the goals of this project, and thank Shrija Mishra, Daejin Choi, Sindhu Ernala, Sandeep Soni, Asra Yousuf, and Dong Whi Yoo for their help and feedback.

REFERENCES

[1] Fortune 500. 2019. `fortune.com/fortune500/list`. (2019). Accessed: 2019-03-27.

[2] Faiyaz Al Zamal, Wendy Liu, and Derek Ruths. 2012. Homophily and latent attribute inference: Inferring latent attributes of twitter users from neighbors. In *ICWSM*.

[3] Noora Albalooshi, Nikolaos Mavridis, and Nabeel Al-Qirim. 2012. A survey on social networks and organization development. In *2012 International Conference on Collaboration Technologies and Systems (CTS)*. IEEE, 539–545.

[4] Mats Alvesson and Stefan Sveningsson. 2015. *Changing organizational culture: Cultural change work in progress*. Routledge.

[5] Aris Anagnostopoulos, Ravi Kumar, and Mohammad Mahdian. 2008. Influence and correlation in social networks. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 7–15.

[6] A Archambault and J Grudin. 2012. A Longitudinal Study of Facebook, LinkedIn, and Twitter Use,âĂŰ in Proceedings of CHI 2012. (2012).

[7] Ghodratollah Bagheri, Reihaneh Zarei, and Mojtaba Nik Aeen. 2012. Organizational silence (basic concepts and its development factors). *Ideal Type of Management* (2012).

[8] Arnold B Bakker, Evangelia Demerouti, and L Lieke. 2012. Work engagement, performance, and active learning: The role of conscientiousness. *Journal of vocational behavior* 80, 2 (2012), 555–564.

[9] Artur Baldauf, Heribert Reisinger, and William C Moncrief. 1999. Examining motivations to refuse in industrial mail surveys. *Journal of the Market Research Society* (1999).

[10] Timothy Baldwin, Paul Cook, Marco Lui, Andrew MacKinlay, and Li Wang. 2013. How noisy social media text, how diffrnt social media sources?. In *Proceedings of the Sixth International Joint Conference on Natural Language Processing*. 356–364.

[11] Stephen R Barley, Gordon W Meyer, and Debra C Gash. 1988. Cultures of culture: Academics, practitioners and the pragmatics of normative control. *Administrative science quarterly* (1988), 24–60.

[12] Murray R Barrick and Michael K Mount. 1991. The big five personality dimensions and job performance: a meta-analysis. *Personnel psychology* 44, 1 (1991), 1–26.

[13] Yehuda Baruch and Brooks C Holtom. 2008. Survey response rate levels and trends in organizational research. *Human relations* 61, 8 (2008), 1139–1160.

[14] Victor Battistich and Allen Hom. 1997. The relationship between students' sense of their school as a community and their involvement in problem behaviors. *American journal of public health* (1997).

[15] Simon J Bell and Bulent Menguc. 2002. The employee-organization relationship, organizational citizenship behaviors, and superior service quality. *Journal of retailing* 78, 2 (2002), 131–146.

[16] Bennett M Berger. 1995. *An essay on culture: Symbolic structure and social structure*. Univ of California Press.

[17] Robert J Bies. 1989. Organizational citizenship behavior: The good soldier syndrome. (1989).

[18] Danah M Boyd and Nicole B Ellison. 2007. Social network sites: Definition, history, and scholarship. *Journal of computer-mediated Communication* 13, 1 (2007), 210–230.

[19] Daniel M Cable and Timothy A Judge. 1996. Person–organization fit, job choice decisions, and organizational entry. *Organizational behavior and human decision processes* 67, 3 (1996), 294–311.

[20] Michael A Campion, Ellen M Papper, and Gina J Medsker. 1996. Relations between work team characteristics and effectiveness: A replication and extension. *Personnel psychology* (1996).

[21] Claire C Caruso, Tim Bushnell, Donald Eggerth, Anneke Heitmann, Bill Kojola, Katharine Newman, Roger R Rosa, Steven L Sauter, and Bryan Vila. 2006. Long working hours, safety, and health: toward a National Research Agenda. *American journal of industrial medicine* 49, 11 (2006), 930–942.

[22] Ralph Catalano. 1979. *Health, behavior and the community: An ecological perspective*. Pergamon Press New York.

[23] Luigi Luca Cavalli-Sforza and Marcus W Feldman. 1981. *Cultural transmission and evolution: A quantitative approach*. Number 16. Princeton University Press.

[24] Alexandra Cetto, Julia Klier, Mathias Klier, Alexander Richter, and Katharina Wiesneth. 2016. The Blessing of Giving: Kowledge Sharing and Knowledge seeking in Enterprise Social Networks.. In *ECIS*. ResearchPaper64.

[25] A Chamberlain. 2015. *Does company culture pay off? Analyzing stock performance of'Best Places to Work'companies*. Technical Report. Glassdoor Research Report, March 2015.

[26] A Chamberlain. 2017. *Give to Get: A Mechanism to Reduce Bias in Online Reviews*. Technical Report. Glassdoor Research Report, October 2017.

[27] Christy MK Cheung, Matthew KO Lee, and Zach WY Lee. 2013. Understanding the continuance intention of knowledge sharing in online communities of practice through the post-knowledge-sharing evaluation processes. *Journal of the American Society for Information Science and Technology* (2013).

[28] Fortune 100 Best Companies. 2019. `fortune.com/best-companies/list/`. (2019). Accessed: 2019-03-27.

[29] Robert A Cooke and Denise M Rousseau. 1988. Behavioral norms and expectations: A quantitative approach to the assessment of organizational culture. *Group & Organization Studies* 13, 3 (1988), 245–273.

[30] Robert A Cooke and Janet L Szumal. 1993. Measuring normative beliefs and shared behavioral expectations in organizations: The reliability and validity of the Organizational Culture Inventory. *Psychological reports* 72, 3_suppl (1993), 1299–1330.

[31] Robert A Cooke and Janet L Szumal. 2000. Using the organizational culture inventory to understand the operating cultures of organizations. *Handbook of organizational culture and climate* 4 (2000), 1032–1045.

[32] Jacqueline A-M Coyle-Shapiro. 2002. A psychological contract perspective on organizational citizenship behavior. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior* 23, 8 (2002), 927–946.

[33] Linda L Dahlberg and Etienne G Krug. 2006. Violence: a global public health problem. *Ciência & Saúde Coletiva* 11 (2006), 1163–1178.

[34] Vedant Das Swain, Koustuv Saha, Hemang Rajvanshy, Anusha Sirigiri, Julie M. Gregg, Suwen Lin, Gonzalo J. Martinez, Stephen M. Mattingly, Shayan Mirjafari, Raghu Mulukutla, and et al. 2019. A Multisensor Person-Centered Approach to Understand the Role of Daily Activities in Job Performance with Organizational Personas. *Proc. ACM IMWUT* (2019).

[35] Munmun De Choudhury and Scott Counts. 2013. Understanding affect in the workplace via social media. In *Proceedings of the 2013 conference on Computer supported cooperative work*. ACM, 303–316.

[36] Daniel R Denison, Stephanie Haaland, and Paulo Goelzer. 2004. Corporate culture and organizational effectiveness: Is Asia different from the rest of the world? *Organizational dynamics* 33, 1 (2004), 98–109.

[37] Rohit Deshpandé, John U Farley, and Frederick E Webster Jr. 1993. Corporate culture, customer orientation, and innovativeness in Japanese firms: a quadrad analysis. *Journal of marketing* 57, 1 (1993), 23–37.

[38] Joan DiMicco, David R Millen, Werner Geyer, Casey Dugan, Beth Brownholtz, and Michael Muller. 2008. Motivations for social networking at work. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*. ACM, 711–720.

[39] Judith S Donath and others. 1999. Identity and deception in the virtual community. *Communities in cyberspace* 1996 (1999), 29–59.

[40] Toni L Doolen, Marla E Hacker, and Eileen M Van Aken. 2003. The impact of organizational context on work team effectiveness: A study of production team. *IEEE Transactions on Engineering Management* 50, 3 (2003), 285–296.

[41] Gabriel Doyle, Amir Goldberg, Sameer Srivastava, and Michael Frank. 2017. Alignment at work: Using language to distinguish the internalization and self-regulation components of cultural fit in organizations. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*.

[42] Toby Marshall Egan, Baiyin Yang, and Kenneth R Bartlett. 2004. The effects of organizational learning culture and job satisfaction on motivation to transfer learning and turnover intention. *Human resource development quarterly* (2004).

[43] Robert Eisenberger, Peter Fasolo, and Valerie Davis-LaMastro. 1990. Perceived organizational support and employee diligence, commitment, and innovation. *Journal of applied psychology* 75, 1 (1990), 51.

[44] Ethan Fast, Binbin Chen, and Michael S Bernstein. 2016. Empath: Understanding topic signals in large-scale text. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 4647–4657.

[45] Robert S Feldman, James A Forrest, and Benjamin R Happ. 2002. Self-presentation and verbal deception: Do self-presenters lie more? *Basic and applied social psychology* 24, 2 (2002), 163–170.

[46] Franco Fiordelisi and Ornella Ricci. 2014. Corporate culture and CEO turnover. *Journal of Corporate Finance* (2014).

[47] Suzy Fox, Paul E Spector, Angeline Goh, Kari Bruursema, and Stacey R Kessler. 2012. The deviant citizen: Measuring potential positive relations between counterproductive work behaviour and organizational citizenship behaviour. *J. Occup. Organ. Psychol* (2012).

[48] Werner Geyer, Casey Dugan, Joan DiMicco, David R Millen, Beth Brownholtz, and Michael Muller. 2008. Use and reuse of shared lists as a social content type. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1545–1554.

[49] Saby Ghoshray. 2013. Employer surveillance versus employee privacy: The new reality of social media and workplace privacy. *N. Ky. L. Rev.* (2013).

[50] Susan R Glaser, Sonia Zamanou, and Kenneth Hacker. 1987. Measuring and interpreting organizational culture. *Management communication quarterly* 1, 2 (1987), 173–198.

[51] Erving Goffman. 1981. *Forms of talk*. University of Pennsylvania Press.

[52] Amir Goldberg, Sameer B Srivastava, V Govind Manian, William Monroe, and Christopher Potts. 2016. Fitting in or standing out? The tradeoffs of structural and cultural embeddedness. *American Sociological Review* (2016).

[53] Abbas Golestani, Mikhil Masli, N Sadat Shami, Jennifer Jones, Abhilash Menon, and Joydeep Mondal. 2018. Real-Time Prediction of Employee Engagement Using Social Media and Text Mining. In *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*. IEEE, 1383–1387.

[54] Maarten Goos, Alan Manning, and Anna Salomons. 2011. Explaining job polarization: the roles of technology, offshoring and institutions. *Offshoring and Institutions (December 1, 2011)* (2011).

[55] Jeffrey H Greenhaus and Saroj Parasuraman. 1993. Job performance attributions and career advancement prospects: An examination of gender and race effects. *Organizational Behavior and Human Decision Processes* 55, 2 (1993), 273–297.

[56] Ulrike Grömping. 2009. Variable importance assessment in regression: linear regression versus random forest. *The American Statistician* (2009).

[57] Glassdoor Community Guidelines. 2019. `https://https://help.glassdoor.com/article/Community-Guidelines/en_US`. (2019). Accessed: 2019-03-21.

[58] John Joseph Gumperz and Dell H Hymes. 1986. *Directions in sociolinguistics: The ethnography of communication*. B. Blackwell.

[59] Ido Guy, Inbal Ronen, Naama Zwerdling, Irena Zuyev-Grabovitch, and Michal Jacovi. 2016. What is Your Organization'Like'?: A Study of Liking Activity in the Enterprise. In *Proc. CHI*.

[60] Heike Heidemeier and Klaus Moser. 2009. Self–other agreement in job performance ratings: A meta-analytic test of a process model. *Journal of Applied Psychology* (2009).

[61] Geert Hofstede. 2001. *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*. Sage publications.

[62] Geert Hofstede, Bram Neuijen, Denise Daval Ohayv, and Geert Sanders. 1990. Measuring organizational cultures: A qualitative and quantitative study across twenty cases. *Administrative science quarterly* (1990).

[63] Raimondo Ingrassia. 2017. Appearance or reality? Monitoring of employer branding in public network space: the Glassdoor case. (2017).

[64] Willow S Jacobson and Shannon Howle Tufts. 2013. To post or not to post: Employee rights and social media. *Review of public personnel administration* (2013).

[65] Karen A Jehn. 1997. A qualitative analysis of conflict types and dimensions in organizational groups. *Administrative science quarterly* (1997), 530–557.

[66] Jeffrey V Johnson and Jane Lipscomb. 2006. Long working hours, occupational health and the changing nature of work organization. *American journal of industrial medicine* 49, 11 (2006), 921–929.

[67] Baek-Kyoo Joo. 2010. Organizational commitment for knowledge workers: The roles of perceived organizational learning culture, leader–member exchange quality, and turnover intention. *Human resource development quarterly* (2010).

[68] Ron Kohavi and others. 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai*.

[69] Peter Kollock. 1999. The economies ol online cooperation. *Communities in cyberspace* 220 (1999).

[70] Hui-Min Lai and Tsung Teng Chen. 2014. Knowledge sharing in interest online communities: A comparison of posters and lurkers. *Comput. Hum. Behav* (2014).

[71] Chei Sian Lee and Long Ma. 2012. News sharing in social media: The effect of gratifications and prior experience. *Computers in human behavior* 28, 2 (2012), 331–339.

[72] Jongseo Lee and Juyoung Kang. 2017. A Study on Job Satisfaction Factors in Retention and Turnover Groups using Dominance Analysis and LDA Topic Modeling with Employee Reviews on Glassdoor. com. (2017).

[73] Jeffrey A LePine, Jason A Colquitt, and Amir Erez. 2000. Adaptability to changing task contexts: Effects of general cognitive ability, conscientiousness, and openness to experience. *Personnel psychology* 53, 3 (2000), 563–593.

[74] Michael Luca and Georgios Zervas. 2016. Fake it till you make it: Reputation, competition, and Yelp review fraud. *Management Science* 62, 12 (2016), 3412–3427.

[75] Ning Luo, Yilu Zhou, and John Shon. 2016. Employee satisfaction and corporate performance: Mining employee reviews on glassdoor. com. (2016).

[76] Gloria Mark, Shamsi T Iqbal, Mary Czerwinski, Paul Johns, Akane Sano, and Yuliya Lutchyn. 2016. Email duration, batching and self-interruption: Patterns of email use on productivity and stress. In *Proc. CHI*.

[77] Stephen M. Mattingly, Julie M. Gregg, Pino Audia, Ayse Elvan Bayraktaroglu, Andrew T. Campbell, Nitesh V. Chawla, Vedant Das Swain, Munmun De Choudhury, Sidney K. D'Mello, Anind K. Dey, and et al. 2019. The Tesserae Project: Large-Scale, Longitudinal, In Situ, Multimodal Sensing of Information Workers. In *CHI Ext. Abstracts*.

[78] Shayan Mirjafari, Kizito Masaba, Ted Grover, Weichen Wang, Pino Audia, Andrew T. Campbell, Nitesh V. Chawla, Vedant Das Swain, Munmun De Choudhury, Anind K. Dey, and et al. 2019. Differentiating Higher and Lower Job Performers in the Workplace Using Mobile Sensing. *Proc. ACM IMWUT* (2019).

[79] Tanushree Mitra, Michael Muller, N Sadat Shami, Abbas Golestani, and Mikhil Masli. 2017. Spread of Employee Engagement in a Large Organizational Network: A Longitudinal Analysis. *PACM HCI* 1, CSCW (2017), 81.

[80] Andy Moniz. 2016. Inferring employeesâĂŹ social media perceptions of goal-setting corporate cultures and the link to firm value. (2016).

[81] Elizabeth Wolef Morrison and Frances J Milliken. 2000. Organizational silence: A barrier to change and development in a pluralistic world. *Academy of Management review* (2000).

[82] Stephan J Motowidlo. 2003. Job performance. *Handbook of psychology: Industrial and organizational psychology* 12 (2003), 39–53.

[83] Arjun Mukherjee, Vivek Venkataraman, Bing Liu, and Natalie Glance. 2013. What yelp fake review filter might be doing?. In *ICWSM*.

[84] Michael Muller, N Sadat Shami, Shion Guha, Mikhil Masli, Werner Geyer, and Alan Wild. 2016. Influences of peers, friends, and managers on employee engagement. In *Proc. GROUP*.

[85] Duc Duy Nguyen, Linh Nguyen, and Vathunyoo Sila. 2019. Does Corporate Culture Affect Bank Risk-Taking? Evidence from Loan-Level Data. *British Journal of Management* 30, 1 (2019), 106–133.

[86] Glassdoor Terms of Use. 2019. `https://www.glassdoor.com/about/terms.htm`. (2019). Accessed: 2019-03-28.

[87] Alexandra Olteanu, Carlos Castillo, Fernando Diaz, and Emre Kiciman. 2016. Social data: Biases, methodological pitfalls, and ethical boundaries. (2016).

[88] O*Net. 2019. About this Site`https://www.glassdoor.com/about/terms.htm`. (2019). Accessed: 2019-03-28.

[89] Charles O'Reilly. 1989. Corporations, culture, and commitment: Motivation and social control in organizations. *California management review* (1989).

[90] Charles A O'Reilly III, Jennifer Chatman, and David F Caldwell. 1991. People and organizational culture: A profile comparison approach to assessing person-organization fit. *Academy of management journal* (1991).

[91] Dennis W Organ. 1997. Organizational citizenship behavior: It's construct clean-up time. *Human performance* 10, 2 (1997), 85–97.

[92] Udai Narain Pareek. 1997. *Training instruments for human resource development*. Tata McGraw-Hill.

[93] Stefan Pasch. 2018. Corporate Culture and Industry-Fit: A Text Mining Approach. (2018).

[94] Jeffrey Pennington, Richard Socher, and Christopher Manning. 2014. Glove: Global vectors for word representation. In *Proc. EMNLP*.

[95] Vinodkumar Prabhakaran, Huzaifa Neralwala, Owen Rambow, and Mona T Diab. 2012. Annotations for Power Relations on Email Threads.. In *LREC*. 806–811.

[96] Daniel I Prajogo and Christopher M McDermott. 2011. The relationship between multidimensional organizational culture and performance. *International Journal of Operations & Production Management* (2011).

[97] Robert E Quinn and John Rohrbaugh. 1983. A spatial model of effectiveness criteria: Towards a competing values approach to organizational analysis. *Management science* 29, 3 (1983), 363–377.

[98] Miguel A Quińones, J Kevin Ford, and Mark S Teachout. 1995. The relationship between work experience and job performance: A conceptual and meta-analytic review. *Personnel psychology* 48, 4 (1995), 887–910.

[99] Glassdoor Fraudulent Reviews. 2019. `help.glassdoor.com/article/Fraudulent-reviews/en_US/`. (2019). Accessed: 2019-04-03.

[100] D Rousseau. 1990. Quantitative assessment of organizational culture: The case for multiple. (1990).

[101] Koustuv Saha, Ayse E. Bayraktaroglu, Andrew T. Campbell, Nitesh V. Chawla, Munmun De Choudhury, Sidney K. D'Mello, Anind K. Dey, Ge Gao, Julie M. Gregg, Krithika Jagannath, and et al. 2019a. Social Media as a Passive Sensor in Longitudinal Studies of Human Behavior and Wellbeing. In *CHI Ext. Abstracts*.

[102] Koustuv Saha, Manikanta D Reddy, Vedant das Swain, Julie M Gregg, Ted Grover, Suwen Lin, Gonzalo J Martinez, Stephen M Mattingly, Shayan Mirjafari, Raghu Mulukutla, and others. 2019c. Imputing Missing Social Media Data Stream in Multisensor Studies of Human Behavior. In *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE.

[103] Koustuv Saha, Manikanta D Reddy, and Munmun De Choudhury. 2019b. JobLex: A Lexico-Semantic Knowledgebase of Occupational Information Descriptors. In *SocInfo*.

[104] Koustuv Saha, Manikanta D Reddy, Stephen M Mattingly, Edward Moskal, Anusha Sirigiri, and Munmun De Choudhury. 2019d. LibRA: On LinkedIn based Role Ambiguity and Its Relationship with Wellbeing and Job Performance. *PACM HCI* CSCW (2019).

[105] Koustuv Saha, Ingmar Weber, and Munmun De Choudhury. 2018. A Social Media Based Examination of the Effects of Counseling Recommendations After Student Deaths on College Campuses. In *ICWSM*.

[106] Jesus F Salgado. 1997. The five factor model of personality and job performance in the European Community. *Journal of Applied psychology* (1997).

[107] Jae San Park and Tae Hyun Kim. 2009. Do types of organizational culture matter in nurse job satisfaction and turnover intention? *Leadership in Health Services* (2009).

[108] Frank L Schmidt, John E Hunter, and Alice N Outerbridge. 1986. Impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance. *Journal of applied psychology* (1986).

[109] Neal Schmitt, Jose M Cortina, Michael J Ingerick, and Darin Wiechmann. 2003. Personnel selection and employee performance. *Handbook of psychology: Industrial and organizational psychology* (2003).

[110] N Sadat Shami, Michael Muller, Aditya Pal, Mikhil Masli, and Werner Geyer. 2015. Inferring employee engagement from social media. In *Proc. CHI*.

[111] N Sadat Shami, Jiang Yang, Laura Panc, Casey Dugan, Tristan Ratchford, Jamie C Rasmussen, Yannick M Assogba, Tal Steier, Todd Soule, Stela Lupushor, and others. 2014. Understanding employee social media chatter with enterprise social pulse. In *Proc. CSCW*.

[112] Walter C Shipley. 2009. *Shipley-2: manual*. WPS.

[113] Joanne Silvester, Neil R Anderson, and Fiona Patterson. 1999. Organizational culture change: An inter-group attributional analysis. *J. Occup. Organ. Psychol* (1999).

[114] Meredith M Skeels and Jonathan Grudin. 2009. When social networks cross boundaries: a case study of workplace use of facebook and linkedin. In *GROUP*.

[115] Linda Smircich. 1983. Concepts of culture and organizational analysis. *Administrative science quarterly* (1983), 339–358.

[116] Jasper Snoek, Hugo Larochelle, and Ryan P Adams. 2012. Practical bayesian optimization of machine learning algorithms. In *Advances in neural information processing systems*. 2951–2959.

[117] Christopher J Soto and Oliver P John. 2017. The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *J. Pers. Soc. Psychol* (2017).

[118] Kate Sparks, Cary Cooper, Yitzhak Fried, and Arie Shirom. 1997. The effects of hours of work on health: a meta-analytic review. *Journal of occupational and organizational psychology* 70, 4 (1997), 391–408.

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

[120] Glassdoor Statistics. 2018. `expandedramblings.com/index.php/numbers-15-interesting-glassdoor-statistics`. (2018). Accessed: 2019-03-22.

[121] Julian Haynes Steward. 1972. *Theory of culture change: The methodology of multilinear evolution*. University of Illinois Press.

[122] Rainer Strack, Carsten Von Der Linden, Mike Booker, and Andrea Strohmayr. 2014. Decoding global talent. *BCG Perspectives* (2014).

[123] Prasanna Tambe and Lorin M Hitt. 2012. Now IT's personal: Offshoring and the shifting skill composition of the US information technology workforce. *Management Science* 58, 4 (2012), 678–695.

[124] Bill Tancer. 2014. *Everyone's a critic: Winning customers in a review-driven world*. Penguin.

[125] Robert P Tett, Douglas N Jackson, and Mitchell Rothstein. 1991. Personality measures as predictors of job performance: A meta-analytic review. *Personnel psychology* 44, 4 (1991), 703–742.

[126] Jennifer Thom and David R Millen. 2012. Stuff IBMers Say: Microblogs as an Expression of Organizational Culture. In *ICWSM*.

[127] Jennifer Thom-Santelli, David R Millen, and Darren Gergle. 2011. Organizational acculturation and social networking. In *Proc. CSCW*.

[128] Glassdoor Give to Get Policy. 2018. `help.glassdoor.com/article/Give-to-get-policy/en_US`. (2018). Accessed: 2019-04-04.

[129] Thea F Van de Mortel and others. 2008. Faking it: social desirability response bias in self-report research. *Australian Journal of Advanced Nursing, The* (2008).

[130] Linn Van Dyne, Ellen Kossek, and Sharon Lobel. 2007. Less need to be there: Cross-level effects of work practices that support work-life flexibility and enhance group processes and group-level OCB. *Human Relations* (2007).

[131] Linn Van Dyne, Don Vandewalle, Tatiana Kostova, Michael E Latham, and LL Cummings. 2000. Collectivism, propensity to trust and self-esteem as predictors of organizational citizenship in a non-work setting. *Journal of organizational behavior* (2000).

[132] James Van Scotter, Stephan J Motowidlo, and Thomas C Cross. 2000. Effects of task performance and contextual performance on systemic rewards. *Journal of applied psychology* 85, 4 (2000), 526.

[133] Yoav Vardi. 2001. The effects of organizational and ethical climates on misconduct at work. *Journal of Business ethics* 29, 4 (2001), 325–337.

[134] Aki Vehtari, Andrew Gelman, and Jonah Gabry. 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing* 27, 5 (2017), 1413–1432.

[135] Brian Westfall. 2019. `softwareadvice.com/resources/job-seekers-use-glassdoor-reviews/`. (2019). Accessed: 2019-06-25.

[136] Larry J Williams and Stella E Anderson. 1991. Job satisfaction and organizational commitment as predictors of organizational citizenship and in-role behaviors. *Journal of management* (1991).

[137] Tzu-Tsung Wong. 2015. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recognition* (2015).

[138] wsj. 2019. Real People. Real Reviews. Real Extortion Scheme?, `https://blogs.wsj.com/law/2010/02/26/real-people-real-reviews-real-extortion-scheme/`. (2019). Accessed: 2019-03-28.

[139] Anbang Xu, Jacob Biehl, Eleanor Rieffel, Thea Turner, and William van Melle. 2012. Learning how to feel again: towards affective workplace presence and communication technologies. In *Proc. CHI*.

[140] Anbang Xu, Haibin Liu, Liang Gou, Rama Akkiraju, Jalal Mahmud, Vibha Sinha, Yuheng Hu, and Mu Qiao. 2016. Predicting perceived brand personality with social media. In *ICWSM*.

[141] Cengiz Yilmaz and Ercan Ergun. 2008. Organizational culture and firm effectiveness: An examination of relative effects of culture traits and the balanced culture hypothesis in an emerging economy. *Journal of world business* 43, 3 (2008), 290–306.

[142] Hui Zhang, Munmun De Choudhury, and Jonathan Grudin. 2014. Creepy but inevitable?: the evolution of social networking. In *Proc. CSCW*. ACM.

[143] Dejin Zhao and Mary Beth Rosson. 2009. How and why people Twitter: the role that micro-blogging plays in informal communication at work. In *Proc. GROUP*.