IDENTIFYING AND QUANTIFYING CULTURAL FACTORS THAT MATTER TO THE IT WORKFORCE: AN APPROACH BASED ON A....

Conference Paper · June 2016

CITATIONS
2

READS
104

4 authors, including:

Theresa Schmiedel
University of Liechtenstein
31 PUBLICATIONS **388** CITATIONS
SEE PROFILE

Oliver Müller
IT University of Copenhagen
79 PUBLICATIONS **625** CITATIONS
SEE PROFILE

Jan vom Brocke
University of Liechtenstein
407 PUBLICATIONS **3,841** CITATIONS
SEE PROFILE

IDENTIFYING AND QUANTIFYING CULTURAL FACTORS THAT MATTER TO THE IT WORKFORCE: AN APPROACH BASED ON AUTOMATED CONTENT ANALYSIS

Research

Schmiedel, Theresa, University of Liechtenstein, Vaduz, Principality of Liechtenstein, theresa.schmiedel@uni.li

Müller, Oliver, IT University of Copenhagen, Copenhagen, Denmark, oliver.mueller@itu.dk

Debortoli, Stefan, University of Liechtenstein, Vaduz, Principality of Liechtenstein, stefan.debortoli@uni.li

vom Brocke, Jan, University of Liechtenstein, Vaduz, Principality of Liechtenstein, jan.vom.brocke@uni.li

Abstract

Organizational culture represents a key success factor in highly competitive environments, such as, the IT sector. Thus, IT companies need to understand what makes up a culture that fosters employee performance. While existing research typically uses self-report questionnaires to study the relation of culture and the success of companies, the validity of this approach is often discussed and researchers call for new ways of studying culture. Therefore, our research goal is to present an alternative approach to culture analysis for examining which cultural factors matter to the IT workforce. Our study builds on 112,610 online reviews of Fortune 500 IT companies collected from Glassdoor, an online platform on which current and former employees can anonymously review companies and their management. We perform an automated content analysis to identify cultural factors that employees emphasize in their reviews. Through a regression analysis on numerical employee satisfaction ratings, we find that a culture of learning and performance orientation contributes to employee motivation, while a culture of assertiveness and gender inegalitarianism has a strong negative influence on employees' satisfaction in the IT workforce. Future research can apply our approach as an alternative method to quantifying culture and its impact on other variables.

Keywords: culture, IT workforce, company reviews, automated content analysis.

1 Introduction

Culture plays a key role in succeeding in highly competitive environments, such as, the IT sector (Egan, Yang, & Bartlett, 2004; Weisinger & Trauth, 2003). Particularly organizational culture has a significant influence on the performance of companies (Chamberlain, 2015; Wilkins & Ouchi, 1983). For decades, managers and researchers have been concerned with organizational culture, as many practitioners experience challenges and benefits that come along with a particular corporate culture (Schein, 2004). However, company culture remains an elusive phenomenon that is hard to grasp (L. C. Harris & Ogbonna, 2002). But to actively manage culture, IT companies need to know what makes up a culture that fosters employee satisfaction in the IT sector and, thus, supports corporate success.

In studying what type of culture influences the performance of organizations, research typically uses self-report questionnaires to quantify culture and related concepts (Taras, Rowney, & Steel, 2009), for example, to examine the impact of culture on the development, implementation, or management of IT (Leidner & Kayworth, 2006). While existing survey instruments are repeatedly used to study culture phenomena, researchers are aware that self-report questionnaires may not provide valid data (Schein, 1990) and, thus, call for new ways of studying the intangible phenomenon (Taras, et al., 2009).

Against this background, the goal of our research is to present an alternative quantitative approach to studying culture and to demonstrate its utility by identifying cultural factors that matter to the IT workforce. To address this objective, we examine online reviews that employees wrote about companies they worked with. Our study builds on 112,610 reviews of Fortune 500 IT companies from the online platform Glassdoor. We perform an automated content analysis, using probabilistic topic modeling, to identify distinct cultural factors that are relevant to employees. Subsequently, we use regression analysis to quantify the relationships these factors have on employee satisfaction.

Next, our paper provides background on culture research in general and on culture in IT research in particular. We then outline the methods we used, specifying automated content analysis and regression analysis. Subsequently, we present the cultural factors we identified to be relevant to employees, quantify their influence on satisfaction of the IT workforce, and compare them to an established culture framework. We then discuss implications and limitations of our study before we conclude the paper.

2 Background

2.1 Culture research

Culture refers to a group's shared values, that is, its basic assumptions about what is right and wrong to do (Schein, 2004). While various cultural layers are important in an organizational context, such as national, organizational, departmental, or professional cultures (Karahanna, Evaristo, & Srite, 2005; Straub, Loch, Evaristo, Karahanna, & Srite, 2002), our study focuses on organizational culture, that is, the patterns of behavior and structure that employees learn over time as the correct way to solve problems in a particular company (Schein, 1990).

Research typically examines group cultures, such as organizations or nations, based on existing value dimensions (Hall & Hall, 1990; Hofstede, 2005; Schwarz, 1999). In fact, a broad range of frameworks exists that distinguish such culture dimensions (Jung et al., 2009; Taras, et al., 2009). For example, organizational culture research often uses culture models, such as the Competing Values Framework (Quinn & Rohrbaugh, 1983) or the sociability-solidarity dimensions (Goffee & Jones, 1996). While these two frameworks are well established, they only include two dimensions for culture analysis.

Research that intends to examine a broader set of cultural dimensions often builds on the culture dimensions from the GLOBE study (House, Hanges, Javidan, Dorman, & Gupta, 2004). With more than 17,000 participants, the study represents one of the most comprehensive culture studies that have been conducted to date. It collects data on pre-defined culture dimensions from 951 organizations in 62 so-

cieties and focuses on the financial, food processing, and telecommunications industries. The dimensions from the GLOBE study are applicable to both the organizational and the societal culture level (Jung, et al., 2009). Table 1 provides an overview of the study's nine culture dimensions.

Cultural dimension	Definition
Assertiveness	The degree to which members of a society are assertive, confrontational or aggressive in social relationships
Future orientation	The degree to which individuals in organizations or societies engage in future-orientated behaviors such as planning, investing in the future, and delaying individual or collective gratification
Gender egalitarianism	The degree to which a society minimizes gender role differences while promoting gender equality
Humane orientation	The degree to which members of a society encourage and reward individuals for being fair, altruistic, friendly, generous, caring and kind to others
In-group collectivism	The degree to which individuals express pride, loyalty and cohesiveness in their organizations or families
Institutional collectivism	The degree to which organizational and societal institutional practices encourage and reward the collective distribution of resources and collective action
Performance orientation	The degree to which an organization or society encourages and rewards members for performance improvement and excellence
Power distance	The degree to which members of a society expect and agree that power should be stratified and concentrated at higher levels of an organization or government
Uncertainty avoidance	The extent to which members of a society seek certainty in their environment by relying on established social norms, rituals and bureaucratic practices

Table 1. Nine cultural dimensions (House, Quigley, & de Luque, 2010)

The GLOBE study examines how far these dimensions *are currently* present in organizations and societies, but the findings mainly focus on differences between societies, as these are found to be larger than variations between organizations. Further, the study examines how far the culture dimensions *should be* present. While the origin of the differences between as-is and to-be scores of the culture dimensions is largely discussed (e.g., Brewer & Venaik, 2010), literature on the GLOBE study does not focus on culture management that works towards desired cultural to-be states in organizations.

2.2 Culture in IT-related research

Studies on organizational culture cover a wide range of industries. While IT-related research has focused on IT-specific culture phenomena, such as the concept of IT culture, which refers to the values a group attributes to IT (Leidner & Kayworth, 2006), we observe a lack of research that focuses on what makes up a successful organizational culture in the IT sector. We believe that understanding cultural factors that support the performance of IT companies is highly relevant because the IT sector today has a strong influence on general economic performance as it affects most if not all industries. Therefore, our study focuses on specifying organizational culture factors that are relevant in the IT industry.

In an IT context, culture is often examined based on self-report questionnaires which assess predefined culture dimensions to analyze how far culture influences IT variables, such as IT design, IT implementation, or IT management (Leidner & Kayworth, 2006). For example, Iivari and Huisman (2007) study the effects of culture on information system development; and Ruppel and Harrington (2001) explore how far organizational culture influences intranet implementation; both studies use a survey approach based on the culture dimensions of the Competing Values Framework (Quinn & Rohrbaugh, 1983).

Another typical stream of research on the role of culture in IT contexts uses interpretive studies to examine the relation between organizational culture and IT variables (Leidner & Kayworth, 2006). For

example, Huang, Newell, Galliers, and Pan (2003) explore how culture influences IT adoption; and Doherty and Doig (2003) analyze how far IT can cause cultural change; both studies use qualitative case analyses and apply existing organizational culture frameworks (Leidner & Kayworth, 2006).

The examples show that IT-related research often studies culture based on existing culture frameworks that provide definitions and measurements of culture dimensions. Such a deductive approach may miss out on identifying novel cultural factors that may be important in an IT context but are not part of existing frameworks, which often stem from outside of the IT field. Such discoveries would be fundamental for culture management, which aims at actively developing culture to increase performance.

The above outlined methodological drawbacks of existing approaches to studying culture may be due to the complexities of studying a large number of organizational cultures via traditional instruments like observations, case studies, or surveys. While observations and case studies allow for inductive identification of cultural characteristics, but are typically restricted to small empirical contexts, surveys can be used in large-scale studies, but typically evaluate the presence of given cultural dimensions in a deductive approach. Our research aims at addressing the call for alternative approaches to culture assessment (Taras, et al., 2009), introducing new ways to study culture inductively on a broad empirical basis via automated content analysis.

3 Methodological Approach

To understand which cultural factors matter to the IT workforce, we follow a two-step process. First, we perform an automated content analysis of the textual parts of employee company reviews to identify topics that represent cultural factors relevant to employees. Second, we use regression analysis to associate the identified topics with numerical company ratings that indicate employee satisfaction.

3.1 Automated content analysis

Our study builds on company reviews from Glassdoor, an online platform on which employees can anonymously review companies and their management. At the time of writing, Glassdoor has a database of more than 8 million employee company reviews from diverse geographical regions and industries. For this study, we focus on reviews of Fortune 500 IT companies, which amount to 112,610 reviews of about 40 companies, spanning a time frame of more than 7 years. In total, the reviews contain about 3.6 million words (i.e., 32 words per review), of which 2,525 are unique words. The mean company rating is 3.41 (Median: 4, Standard Deviation: 1.14), on a scale from 1 (worst) to 5 (best).

To automatically analyze the content of the reviews we use probabilistic topic modeling, in particular Latent Dirichlet Allocation (LDA). LDA is an unsupervised machine learning technique, which is able to inductively identify common topics running through large collections of texts and cluster documents into multiple topical categories (Blei, 2012; Blei, Ng, & Jordan, 2003). The idea underlying LDA, and its predecessor Latent Semantic Analysis (LSA) (Landauer, Foltz, & Laham, 1998), is based on the distributional hypothesis of linguistics (Firth, 1957; Z. Harris, 1954), which states that "words that occur in the same contexts tend to have similar meanings" (Turney & Pantel, 2010, p. 142). For example, if in a sub-group of newspaper articles the words "island", "tourism", "resort", and "hotel" tend to highly co-occur, they are indicative of a common topic, namely "vacation". In statistical terms, LDA represents documents as probability distributions over a vector of topics, and each topic, in turn, is represented as a probability distribution over a controlled vocabulary of words (Figure 1).

LDA requires two main inputs. First, a collection of preprocessed documents. The goal of the preprocessing is to remove as much noise from the documents as possible, hence, the type and amount of preprocessing performed can have a substantial influence on the quality of the resulting topic model. Up to date, no clear standards for preprocessing have emerged and researchers are advised to try different preprocessing strategies in order to yield the best results (Boyd-Graber, Mimno, & Newman, 2014). In our analysis, we removed standard English stopwords (e.g., "the", "and") as well as the

names and abbreviations of the 40 IT companies being part of our study from the review texts. In addition, we limited the analysis to nouns, verbs, adjectives, and adverbs only, disregarding function words bearing little semantic information, such as prepositions, pronouns, or conjunctions. Finally, all remaining words were lemmatized, that is, reduced to their common stems (e.g., turn plural nouns to singular or removing the third person "-s" for verbs). The required second input is the number of topics to be extracted. Again, no best practices for determining an optimal number of topics are available yet (Boyd-Graber, et al., 2014). If the resulting topic model is used in fully automated applications (e.g., search engines, recommender systems), between 200 and 500 topics are typical choices. However, if the results are intended to be interpreted by humans, as it is the case in our study, a smaller number between 20 and 100 topics is more usual. We tried different choices ranging between 20 and 100 topics, and qualitatively evaluated the interpretability of the results. We selected 50 topics as the best choice, as more fine-grained topic models (between 50 and 100 topics) produced near duplicate topics which only differed in their writing style, and more abstract models (between 20 and 50 topics) failed to clearly discriminate between topics by merging two or more similar topics into one.

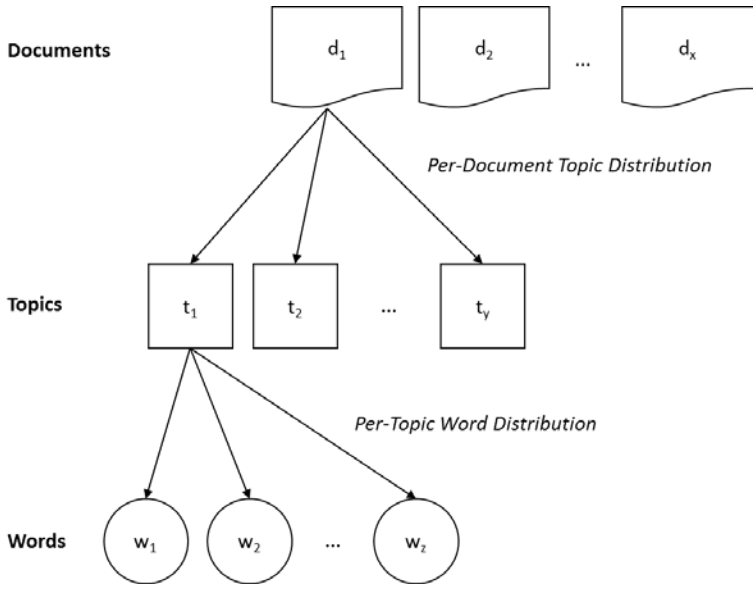

Figure 1. Schematic Overview of LDA

The central output of LDA is the per-topic word distribution, which has the form of a matrix with one row per document (in our case: 112,610) and one column per topic (in our case: 50). The cells of the matrix contain probabilities indicating the prevalence of a topic in a document. Hence, a row of the matrix represents a statistical summary of the content of a document (the probabilities for one document add up to 100%). This data structure, combined with the numerical company rating of each review, served as the input for the second step of our analysis, the regression analysis.

3.2 Regression analysis

The goal of the regression analysis is to quantify relationships between the prevalence of topics in documents (independent variables) and the corresponding numerical company ratings (dependent variable). Following prior research on explaining review star ratings through regression analysis (e.g., Dhar & Chang, 2009; Ganu, Kakodkar, & Marian, 2013), we fit a pooled linear ordinary least square (OLS) model to the data, as specified by the following equation:

Equation (1): $Rating_i = \beta_0 + \beta_{Ti}\, Topics_i + \beta_{Ci}\, Company_i + \varepsilon_i$,

where *i* indexes the *i*th review, β_0 is the regression intercept, and ε_i is the residual random error term. $Rating_i$ is the numerical company rating of a review, $Topics_i$ is an array of topic probabilities derived from the LDA analysis, and β_{Ti} the corresponding array of regression coefficients for each topic. $Company_i$ is a dummy variable representing companies, and β_{Ci} its respective coefficient. Including a dummy for companies allows to control for systematic differences in ratings between companies.

Testing for normal distribution of residuals, homoscedasticity of residuals, multicollinearity of independent variables, and influential cases, revealed no violations of underlying statistical assumptions of the regression model (Field, Jeremy, & Field, 2012). Thus, we proceeded with the interpretation of the analysis results. Next, we report on the findings of our two-step analysis approach.

4 Findings

4.1 Topics in company reviews

Based on our topic modelling approach, we identify 50 distinct topics that employees working in the IT sector repeatedly mention in company reviews. We do not distinguish between employee roles, because we are interested in the general organizational culture rather than subcultures of specific employee types. The 50 topics obviously represent aspects of organizations that are particularly relevant to the IT workforce. In order to interpret and label the topics, three researchers independently coded them based on the most probable words that made up a topic (cf. per-topic word distributions in Figure 1) and based on those reviews that were most closely associated with a topic (cf. per-document topic distributions in Figure 1). For example, Topic 4 is represented by highly likely terms such as "innovative", "idea", "risk", and "creativity" and related reviews referred to an organization that fosters innovation. This topic, which we labelled "innovation", represents a cultural factor relevant to employees. Table 2 provides further examples of topics, their most probable words, exemplary sentences from associated reviews, and the researcher-generated descriptive topic labels. The consolidation of the individual codings revealed an inter-coder agreement of 0.82 percent and an average Kappa value of 0.86 (Light, 1971), which indicates high inter-coder reliability (Moore & Benbasat, 1991). Cases in which the codings differed between researchers were discussed until agreement between the coders was reached. A list of all 50 topics, along with their most probable words, is provided in the appendix.

Topic ID	Most probable terms	Exemplary review	Topic label
4	innovation culture idea company risk innovative focus innovate creativity creative	"Big organization yet fostering innovation and start up mentality."	Innovation
12	woman boy play promote politics club political promotion game diversity	"This place is not a good place for women to work. […] It's still an old boys network."	Gender inequality
21	balance life worklife work good growth great compensation culture personal	"A good place to maintain your work life and personal life balanced."	Work-life balance

Table 2. Exemplary topic labeling

Most identified topics relate to the culture that is present in an organization, that is, they relate to what organizations value and how people interact. Exemplary topics include "great work environment", "fast pace", "short-term focus", "employee treatment", and "teamwork". While we are particularly interested in cultural factors that are relevant to the IT workforce, the topic model analysis also generated topics only describing a special employment status like "internship" (with terms such as "intern", "program", "summer", "mentor", and "student") or employment via a "contractor" (with terms such as "contract", "permanent", "temp", "full", and "hire"), without any hints on the organizational culture.

Some also describe particular work areas such as “call center”, “sales”, “consulting”, “software engineering”, and “retail store”. We excluded these topics from the further analysis and discussion.

4.2 Relevance of cultural factors

In order to examine which of the topics (hereafter called cultural factors or just factors) identified via topic modelling influence company ratings in which direction and to what extent, we performed a regression analysis. Table 3 summarizes the results of this analysis. Overall, our regression model is able to explain a quarter of the variance in company ratings (adjusted R^2: 0.2516). Inspecting the independent variables of the model, we observe that most factors are positively associated with employees’ perception of organizations in the IT sector (27 factors are significantly positively related to employee satisfaction, while 15 factors show a significant negative influence on the dependent variable).

A deeper look at the cultural factors with the largest influence on employees’ evaluations of organizations in the IT sector reveals several insights regarding desired organizational culture in the IT sector.

Most positively perceived cultural factors

- Great work environment:

  The company reviews exhibiting high probability for this topic refer to a friendly atmosphere at the workplace, supporting colleagues, great coworkers, awesome people, interesting and challenging work, and a relaxed but hard working culture. These reviews emphasize the importance of strong interpersonal relationships for motivating employees in their jobs. Beyond, the reviews also reveal that employees value a culture of challenging tasks (e.g., “Fantastic atmosphere, good culture, interesting people, challenging work”). This aspect seems to be related to the factor fast pace.

- Fast pace:

  Company reviews strongly associated with this factor typically refer to a challenging environment, high performance and high pressure, opportunities to learn and grow, fast rates of change, and quick adaptation. The reviews describe that employees in the IT sector appreciate finding their way in a highly dynamic environment that challenges them intellectually (e.g., “The work environment tends to be very stimulating and fast paced.”). They also show that employees value hard work and perceive personal growth as a reward.

- Smart people:

  Typical company reviews associated with this factor refer to dynamic and energetic colleagues, talented people with interesting ideas, intelligent and passionate people that are excited to work on high impact projects which change the world. The reviews show that working together with smart colleagues and learning how to tackle complex and relevant problems very strongly motivates employees (e.g., “Lots of smart, motivated engineers to draw inspiration from.”). The reviews also show that employees work long hours because they are excited to contribute to changing the world.

The three cultural factors above strongly influence the motivation of employees to work with a specific employer in the IT sector. Next, we turn towards aspects of culture that have the strongest negative influence on employee satisfaction in our data.

Most negatively perceived cultural factors

- Bad management

  Company reviews that are highly associated with this factor refer to incompetent and unprofessional management, telling lies, blaming when things go wrong, lazy managers that do not care about employees, and a focus on making money (e.g., “they blame the people rather than work with them to correct the problem”). In the reviews, people repeatedly argue they only stay in the company because of the pay or because it looks good on their resume. The reviews emphasize that a culture of confrontational management styles strongly dissatisfies employees.

- Gender inequality

  Typically, company reviews associated with this factor refer to a lack of women in management positions and a lack of their promotion, old boys networks, discrimination of minorities, lack of general diversity, but also an advantageous bias for women and minorities. The reviews describe that employees in the IT sector are particularly dissatisfied by cultures in which gender role differences prevail in organizations (e.g., "Claims of pro-woman, pro-family is all marketing hype.").

- Senior management

  The most representative company reviews for this factor refer to legacy management, political games, decision-making based on relationships, lack of direct reports to senior executives, and a general gap between senior management and the operational level. The reviews show that a culture of organizational hierarchies dissatisfies employees in the IT sector when there is a lack of direct relation between the hierarchical levels and a lack of perceived contribution of the senior management level (e.g., "Senior management are out of touch with the company.").

Beyond the three examples of negative factors that outline what companies in the IT sector should avoid to foster employee satisfaction and performance, the results show, for example, that the factor benefits is perceived as a neutral factor that, rather surprisingly, does not show an effect on employee satisfaction. Overall, the introduced cultural factors with the largest positive and negative influence on employees' satisfaction provide important insights into what matters to the IT workforce.

Factor	**Estimate**	**Std. Error**	**t value**	**Sign.**
(Intercept)	3.23	0.05	69.82	***
Great work environment	2.14	0.04	52.72	***
Fast pace	1.69	0.06	26.98	***
Smart people	1.45	0.05	32.05	***
Large company	1.38	0.06	22.84	***
Career opportunity	1.25	0.05	26.76	***
Learning	1.25	0.05	26.81	***
Unbureaucratic organization	1.22	0.05	22.90	***
Flexible work hours	1.18	0.05	23.61	***
Traveling	1.15	0.07	16.61	***
Work-life balance	1.08	0.04	27.60	***
High tech	1.08	0.08	14.22	***
Mobility	0.93	0.06	15.20	***
Perks	0.87	0.06	14.50	***
Innovation	0.82	0.05	15.62	***
Internal processes	0.78	0.06	12.52	***
Decision making	0.72	0.07	10.15	***
Salary	0.71	0.04	16.38	***
Time management	0.64	0.06	11.18	***
Individual performance goals	0.60	0.06	10.36	***
Teamwork	0.57	0.07	8.55	***
Market orientation	0.42	0.05	8.12	***
Appraisal	0.33	0.04	7.59	***

Project work	0.33	0.08	4.28	***
Customer service	0.32	0.06	5.20	***
Reorganization	0.28	0.07	4.27	***
Problem solving	0.26	0.08	3.30	***
Talent management	0.22	0.07	3.07	**
Benefits	0.03	0.06	0.44	
Leadership	-0.19	0.06	-3.12	**
Technical skills	-0.24	0.07	-3.42	***
Young professionals	-0.31	0.08	-3.94	***
Working hours	-0.46	0.05	-8.42	***
Short-term focus	-0.48	0.07	-7.05	***
Leaving	-0.67	0.06	-11.34	***
Bonus	-0.74	0.06	-13.30	***
Middle management	-0.81	0.06	-13.46	***
Performance review	-0.91	0.07	-13.65	***
HR policies	-0.97	0.07	-14.09	***
Employee treatment	-0.97	0.05	-18.23	***
Layoffs	-1.67	0.05	-33.61	***
Senior management	-1.79	0.07	-24.77	***
Gender inequality	-1.80	0.07	-24.31	***
Bad management	-2.63	0.05	-51.89	***
R^2 0.2522, adjusted R^2 0.2516, ***				
Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

Table 3. Regression results (ordered by coefficient estimates)

4.3 Comparison to existing research

To examine the results of our study on a more abstract level, we relate to the culture dimensions of the GLOBE study. To connect our findings to these dimensions, we mapped the automatically identified cultural factors to the nine value dimensions from the GLOBE study. Three researchers individually performed the mapping task. In the initial mapping, an inter-coder agreement of 0.68 percent with an average Kappa value of 0.74 was reached (Light, 1971), indicating sufficient inter-coder reliability (Moore & Benbasat, 1991). When mappings differed, the researchers examined the factors in detail based on most probable terms and reviews and discussed the mapping until agreement was reached.

The mapping of the identified factors to the cultural dimensions of the GLOBE study shows that all nine dimensions are relevant to the IT workforce (Table 4). During the coding consolidation process, coders found that some factors, which are mapped to future orientation, could also be grouped into a sub-dimension related to the concept of learning. In particular, this applied to the factors "learning", "talent management", "smart people", and "technical skills". Lifelong learning and continuous skill development are especially relevant in the IT sector. Thus, the GLOBE dimensions might be extended by this dimension to better fit the specifics of this industry.

Examining the importance of the cultural dimensions for employee satisfaction, the number of mapped cultural factors per GLOBE dimension provides a first indication for their relevance. Particularly, factors in the dimensions performance orientation (13 factors) and humane orientation (6 factors) received intensive attention. Considering the percentage of review texts that focus on each of the cultural dimensions (calculated by cummulating the average per-document topic probabilities of the corre-

sponding topics) confirms the high relevance of these two dimensions (Table 4); 29% of the text of all reviews written relates to the performance dimension, and 16% to the humane orientation dimension.

However, comparing the relative attention cultural dimensions receive in company reviews with the magnitude of related regression coefficients shows that culture dimensions that are prevalent do not necessarily have a strong impact on employee satisfaction. For example, gender egalitarianism is seldom mentioned in company reviews (probability of 1.13%), but if it is mentioned, it has a very strong influence on how employees assess their organization (third strongest regression coefficient); performance orientation, in contrast, is extensively discussed in company reviews, but most of the regression coefficient estimates of the corresponding factors are quite small.

Cultural dimension	**Topics identified through automated content analysis**	**Cummulated percentage of review texts**
Assertiveness	Bad management, Layoffs, Leaving	7.67
Future orientation	High tech, Innovation, Reorganization, Short-term focus, Young professionals	7.50
Learning	*Learning, Smart people, Talent management, Technical skills*	*10.07*
Gender egalitarianism	Gender inequality	1.13
Humane orientation	Employee treatment, Flexible work hours, HR policies, Perks, Working hours, Work-life balance	16.23
In-group collectivism	Great work environment	4.68
Institutional collectivism	Teamwork, Travelling	2.98
Performance orientation	Appraisal, Benefits, Bonus, Career opportunity, Customer service, Fast pace, Individual performance goals, Market orientation, Mobility, Performance review, Problem solving, Salary, Time management	28.96
Power distance	Decision making, Leadership, Middle management, Senior management	6.19
Uncertainty avoidance	Internal processes, Large company, Project work, Unbureaucratic organization	7.02

Table 4. Mapping of the identified factors to cultural dimensions

Examining the direction (positive or negative) of regression coefficients by GLOBE cultural dimensions provides further interesting insights.

Regarding the positively associated factors, the culture dimensions future orientation, learning, in-group collectivism, institutional collectivism, performance orientation, and uncertainty avoidance are dominant. Also considering the overall amount of factors that have been mapped to these dimensions, we find that employees particularly appreciate a culture of performance orientation and future orientation/learning. The relevance of these dimensions is further emphasized as they are also represented by two of the culture factors with the highest positive regression coefficients.

Regarding the negatively associated factors, the culture dimensions assertiveness, gender egalitarianism, and power distance are dominant. Taking also the number of factors mapped to these culture dimensions into account, we can conclude that employees in the IT sector specifically dislike a culture with issues arising from power distance. This finding is further underlined by the fact that power distance is represented by one of the culture factors that most negatively contributes to employees' perception of an organization.

Finally, the dimension humane orientation is the only culture dimension that contains the same amount of positively and negatively associated factors. This finding shows that both a strong emphasis and a neglect of employees' well-being receive equal attention among employees.

Our results quantitatively estimate the influence of inductively identified cultural factors on employee satisfaction in the IT sector, and, hence, provide insights that go beyond the GLOBE study results. For example, the GLOBE study does not cover how to work towards the desired to-be state of a particular culture, while our study has direct implications for developing organizational cultures into a preferred direction. Examining the importance of cultural dimensions in organizations with regard to a dependent variable (e.g., employee satisfaction) allows for selecting those culture dimensions with the largest effect on organizational performance to then plan effective interventions.

5 Discussion

Based on the comparison of our findings to existing research, we discuss both content-related and methodological implications of our results for research and also point out limitations of our work.

With regard to content, our research provides insights into cultural factors that are particularly relevant to the IT sector. While a culture of fast pace and hard work may not be perceived positively in all contexts, our study shows that these factors have a strong positive influence on employee satisfaction in the IT workforce. Challenging work tasks and an environment of continuous learning are much more appreciated than performance appraisals or benefits. Another important insight refers to factors that have a negative influence on employee satisfaction. Apart from incompetent management or a gap between senior management and operational level, discrimination of women in the IT workforce is among the factors with the strongest negative influence on employee satisfaction. As a recognized theoretical approach in the IT field, the social construction of gender considers cultural patterns essential in creating gender differences (Harding, 1987; Wilson, 2004). Future research may build on the understanding that culture and gender are both socially shaped, connect our findings to existing studies on gender and IT work, and integrate these findings into further studies of the IT workforce.

While our study confirms the overall relevance of the GLOBE dimensions for managing organizational cultures in the IT sector, it complements the GLOBE study and similar culture frameworks in several ways. First, following an inductive approach we identified more fine-grained and more specific factors than, for example, the nine abstract GLOBE dimensions. Second, the GLOBE study focuses on organizations from the financial, food processing, and telecommunications industries, and does not sufficiently cover the IT sector. Third, the GLOBE study reports how far the culture dimensions are currently present or should be present, while our approach to culture analysis allows quantifying the effect of cultural factors on employee satisfaction as a basis for organizational culture management. Future research may build on our findings in examining how to establish a company culture that considers the identified cultural factors in the IT sector.

Regarding the methodological contribution, our study provides guidance on new forms of culture assessments. While automated content analysis represents an established research method in general, it is new to culture research. One particular advantage of this method compared to established survey or case study methods lies in the nature and amount of data that can be considered. Since online reviews are generated without a research purpose in mind, that is, they are not provoked through questionnaires or interview questions, they may contain less bias in terms of social expectations. Furthermore, culture is not assessed deductively based on predefined scales or constructs, as it is typically done in surveys, but inductively based on what employees felt is most important to say about a company or job. Thus, automated content analysis combines the advantages of an inductive qualitative approach to studying culture with the advantages of measuring variable relations quantitatively and in large scale. Therefore, our study contributes to the call for new ways of studying culture (Taras, et al., 2009). Future research may compare our results with analyses of different text types (e.g., internal documents).

Beyond implications for research, our study also has implications for practice. Practitioners from IT companies can use our findings to reflect how far their organizational culture contains those cultural factors that are particularly influential with regard to employees' satisfaction. Ensuring that none of the highly negatively perceived factors, such as bad management or gender inequality, are present in

the organization would positively influence employee perception. In addition, integrating highly positively perceived factors, such as fast pace, smart people, and a learning environment, would help to establish a surrounding that employees appreciate. We argue that considering these factors will largely benefit companies in increasing employee satisfaction and, in turn, contribute to overall performance. Our findings, thus, can help to steer company cultures into the right direction. Beyond, they also provide a blueprint conducting culture assessments in practice using automated content analysis methods.

Naturally, our study also contains limitations. Since we examine data from IT companies of the Fortune 500 list only, the generalizability of our findings is limited. Future research may complement our results through studying the cultures of IT companies outside the United States of America (USA) and through studying the cultures of companies from other industries. Nevertheless, our findings provide first insights into cultural factors that matter to employees in the IT sector of the USA. Furthermore, the approach of using company reviews that are provided on an online platform may further restrict the generalizability of our findings, because only the opinions of those employees that favor online reviews are included in our data. One may question the legitimacy of the reviews since demographic factors such as age, gender, and employment status may have an effect on the reviews, however, controlling for these factors did not reveal substantially different results. Further research may complement our findings, accounting for the subjective nature of culture, by examining corporate cultures including opinions from all employees and using additional qualitative measures. As our research builds on a large number of company reviews, it is suitable to gain first insights on our study purpose.

6 Conclusion

We study organizational culture as a key determinant of employee satisfaction. Focusing on the IT sector, we examine which cultural factors matter to the IT workforce. We use company reviews of Fortune 500 IT companies from the online platform Glassdoor and apply automated content analysis and regression analysis to identify cultural factors that employees find particularly relevant, and to measure their impact in employee satisfaction. To the best of our knowledge, our study is the first to conduct a systematic analysis of company cultures that builds on automated content analysis. Future research can build on our research both from a content and methodological perspective, and practitioners may use our findings to assess and develop their organizational cultures. Against this background, our research can further nurture the fascination of culture among managers and researchers as it allows to gain insights into which cultural factors matter to the IT workforce.

Appendix

Topic	**Most probable words**
Appraisal	hike onsite good appraisal salary india le join work fresher
Bad management	people bad manager fire management stop leave money lie care
Benefits	stock health benefit insurance medical vacation option match plan dental
Bonus	raise bonus increase salary pay year promotion employee performance low
Call center	call phone center day email train supervisor customer job send
Career opportunity	career growth opportunity path development advancement great train progression limit
Consulting	firm client consult government csc consultant staff associate bah contract
Contractor	contractor contract permanent temp full hire badge vendor worker treat
Customer service	service customer business delivery focus unit client support solution model
Decision making	decision make slow empower people change move process impact layer
Employee treatment	employee care treat company respect management asset customer happy work
Fast pace	fast pace grow environment high challenge culture work dynamic learn

Flexible work hours	flexible work hour home schedule flexibility long office time environment
Gender inequality	woman boy play promote politics club political promotion game diversity
Great work environment	work great place environment good con nice friendly atmosphere culture
High tech	edge technology cut ti lead semiconductor industry work engineer late
HR policies	hr employee policy manager open door management feedback department communication
Individual performance goals	individual employee goal feedback contributor performance achieve reward recognize success
Innovation	innovation culture idea company risk innovative focus innovate creativity creative
Internal processes	process internal structure system tool procedure organization complicate complex streamline
Internship	intern internship program summer mentor event project experience student felt
Large company	large group company wide opportunity business variety size organization range
Layoffs	cost cut layoff employee outsource lay year job reduction save
Leadership	leadership vision lack direction clear strategy leader strategic execute execution
Learning	lot learn people good opportunity work place great thing smart
Leaving	year company leave thing join day find make people time
Market orientation	product market company cloud acquisition software business acquire storage consumer
Middle management	management middle level upper senior layer low mid good manager
Mobility	move job position mobility role opportunity movement upward lateral transfer
Performance review	review performance system rank rat stack curve evaluation manager process
Perks	free food gym office campus cafeteria san perk nice coffee
Problem solving	problem solve issue fix data resolve solution time thing address
Project work	project client resource work good bench skill assign interest assignment
Reorganization	change layoff constant year reorgs reorganization direction restructure frequent past
Retail store	retail store customer discount genius product specialist amaze corporate great
Salary	good salary benefit pay compare company low compensation work average
Sales	sale account rep sell quota commission territory customer deal product
Senior management	director vps vp executive level senior dead mgmt sr management
Short-term focus	term long short focus quarter street wall profit cost financial
Smart people	people smart great work amaze hard interest awesome fun perk
Software engineering	engineer software developer product code design development test feature team
Talent management	talent top retain performer attract reward notch people recognize leave
Teamwork	team manager member depend work good lead group vary mate
Technical skills	technical skill manager engineer knowledge management people technically staff manage
Time management	time work spend make hard people feel job thing lot
Traveling	travel local site train global country access international online office
Unbureaucratic organization	company big large tape red lot small feel move bureaucracy
Working hours	hour day week work weekend overtime night shift time vacation
Work-life balance	balance life worklife work good growth great compensation culture personal
Young professionals	college hire experience graduate grad fresh young school recruit interview

Table 5. Topics resulting from automated content analysis

References

Blei, D. (2012). Probabilistic topic models. *Communications of the ACM, 55*(4), 77-84.

Blei, D., Ng, A., & Jordan, M. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research, 3*(1), 993-1022.

Boyd-Graber, J., Mimno, D., & Newman, D. (2014). Care and feeding of topic models: Problems, diagnostics, and improvements. In E. M. Airoldi, D. Blei, E. A. Erosheva & S. E. Fienberg (Eds.), *Handbook of mixed membership models and their applications* (pp. 3-34). Boca Raton: CRC Press.

Brewer, P., & Venaik, S. (2010). GLOBE practices and values: A case of diminishing marginal utility? *Journal of International Business Studies, 41*, 1316-1324.

Chamberlain, A. (2015). Does company culture pay off? Analyzing stock performance of 'Best Places to Work' companies. *Glassdoor Research Report, March 2015*, 1-18.

Dhar, V., & Chang, E. A. (2009). Does chatter matter? The impact of user-generated content on music sales. *Journal of Interactive Marketing, 23*, 300-307.

Doherty, N. F., & Doig, G. (2003). An analysis of the anticipated cultural impacts of the implementation of data warehouses. *IEEE Transactions on Engineering Management, 50*(1), 78-88.

Egan, T. M., Yang, B., & Bartlett, K. R. (2004). The effects of organizational learning culture and job satisfaction on motivation to transfer learning and turnover intention. *Human Resource Development Quarterly, 15*(3), 279-301.

Field, A., Jeremy, M., & Field, Z. (2012). *Discovering statistics using R*. London, Thousand Oaks, New Delhi: Sage.

Firth, J. R. (1957). A synopsis of linguistic theory 1930-1955 *Studies in Linguistic Analysis* (pp. 1-32). Oxford: Philological Society.

Ganu, G., Kakodkar, Y., & Marian, A. (2013). Improving the quality of predictions using textual information in online user reviews. *Information Systems, 38*, 1-15.

Goffee, R., & Jones, G. (1996). What holds the modern company together? *Harvard Business Review, 74*(6), 133–148.

Hall, E. T., & Hall, M. R. (1990). *Understanding cultural differences*. Boston, MA: Intercultural Press.

Harding, S. G. (1987). *Feminism and methodology: Social science issues*. Bloomington: Indiana University Press.

Harris, L. C., & Ogbonna, E. (2002). The unintended consequences of culture interventions: A study of unexpected outcomes. *British Journal of Management, 13*, 31-49.

Harris, Z. (1954). Distributional structure. *Word, 10*(23), 146-162.

Hofstede, G. (2005). *Cultures and organizations* (2 ed.). New York: McGraw-Hill.

House, R. J., Hanges, P. J., Javidan, M., Dorman, P. W., & Gupta, V. (2004). *Culture, leadership, and organizations: The GLOBE study of 62 societies*. Thousand Oaks, London, New Delhi: Sage.

House, R. J., Quigley, N. R., & de Luque, M. S. (2010). Insights form project GLOBE: Extending global advertising research through a contemporary framework. *International Journal of Advertising, 29*(1), 111-139.

Huang, J. C., Newell, S., Galliers, R., & Pan, S. L. (2003). Dangerous liaisons? Component based development and organizational subcultures. *IEEE Transactions on Engineering Management, 50*(1), 89-99.

Iivari, J., & Huisman, M. (2007). The relationship between organizational culture and the deployment of systems development methodologies. *Management Information Systems Quarterly, 31*(1), 35-58.

Jung, T., Scott, T., Davies, H. T. O., Bower, P., Whalley, D., McNally, R., & Mannion, R. (2009). Instruments for exploring organizational culture: A review of the literature. *Public Administration Review, November|December*, 1087-1096.

Karahanna, E., Evaristo, J. R., & Srite, M. (2005). Levels of culture and individual behavior: An investigative perspective. *Journal of Global Information Management, 13*(2), 1-20.

Landauer, R. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes, 25*(2-3), 259-284.

Leidner, D. E., & Kayworth, T. (2006). A review of culture in information systems research: Toward a theory of information technology culture conflict. *Management Information Systems Quarterly*, *30*(2), 357–399.

Light, R. J. (1971). Measures of response agreement for qualitative data: Some generalizations and alternatives. *Psychological Bulletin*, *76*(5), 365-377.

Moore, G. C., & Benbasat, I. (1991). Development of an instrument to measure the perceptions of adopting an information technology innovation. *Information Systems Research, 2*(3), 192-222.

Quinn, R. E., & Rohrbaugh, J. (1983). A spatial model of effectiveness criteria: Towards a competing values approach to organizational analysis. *Management Science, 29*(3), 363-377.

Ruppel, C. P., & Harrington, S. J. (2001). Sharing knowledge through intranets: A study of organizational culture and intranet implementation. *IEEE Transactions on Professional Communication, 44*(1), 37-52.

Schein, E. H. (1990). Organizational culture. *American Psychologist*, *45*(2), 109-119.

Schein, E. H. (2004). *Organizational culture and leadership* (3 ed.). San Francisco: Jossey-Bass.

Schwarz, S. H. (1999). A theory of cultural values and some implications for work. *Applied Psychology: An International Review, 48*(1), 23-47.

Straub, D., Loch, K., Evaristo, R., Karahanna, E., & Srite, M. (2002). Towards a theory-based measurement of culture. *Journal of Global Information Management, 10*(1), 13–23.

Taras, V., Rowney, J., & Steel, R. (2009). Half a century of measuring culture: Review of approaches, challenges, and limitations based on the analysis of 121 instruments for quantifying culture. *Journal of International Management, 15*, 357-373.

Turney, P., & Pantel, P. (2010). From frequency to meaning: Vector space models of semantics. *Journal of Artificial Intelligence Research, 37*, 141-188.

Weisinger, J. Y., & Trauth, E. M. (2003). The importance of situating culture in cross-cultural IT management. *IEEE Transactions on Engineering Management, 50*(1), 26-30.

Wilkins, A. L., & Ouchi, W. G. (1983). Efficient cultures: Exploring the relationship between culture and organizational performance. *Administrative Science Quarterly*, *28*(3), 468-481.

Wilson, M. (2004). A conceptual framework for studying gender in information systems research. *Journal of Information Technology, 19*(1), 81-92.