2021-09-24

Factors Influencing the Effectiveness of Third-Party Online Organizational Reviews

Morgan, Jenelle Alissa

Morgan, J. A. (2021). Factors Influencing the Effectiveness of Third-Party Online Organizational Reviews (Unpublished master's thesis). University of Calgary, Calgary, AB.

<http://hdl.handle.net/1880/114003>

Downloaded from PRISM Repository, University of Calgary

UNIVERSITY OF CALGARY

Factors Influencing the Effectiveness of Third-Party Online Organizational Reviews

by

Jenelle Alissa Morgan

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER OF SCIENCE

GRADUATE PROGRAM IN PSYCHOLOGY

CALGARY, ALBERTA

SEPTEMBER, 2021

© Jenelle Alissa Morgan 2021

Abstract

Online reviews of organizations through portals such as Glassdoor and Indeed are growing in popularity and have the potential to influence organizational attractiveness. This study examined how the various components of online organizational reviews affect helpfulness ratings. We examined the relationship between the sentiment of the reviews (i.e., the extent to which reviews vary from positive to negative attitudes) and helpfulness ratings, as well as the roles of organizational image, deviation from consensus in attitudes and employee status as moderators of this relationship. Part two of this study evaluated, through automated text analysis using Latent Dirichlet Allocation (LDA) and Structural Topic Modeling, the type of topics discussed in reviews, their influence on helpfulness votes, and how they differ based on employee status. After extracting and analyzing 24,687 reviews on Glassdoor, we found that negative reviews generally received higher helpfulness ratings which was more pronounced when they were provided by former employees and deviated from a consensus in attitudes. The topics revealed through the LDA, addressed both the instrumental and symbolic images of organizations that would facilitate inferring need-supplies and supplementary fit. They also presented differential effects on helpfulness ratings, and uniquely interacted with sentiment. These findings illustrate the impact that negative attitudes can have on helpfulness ratings, and they further highlight the need for organizations to consider the implications of how their brand images are portrayed online.

Keywords:

Online organizational reviews, word-of-mouth, online recruiting, text analysis, helpfulness ratings.

Preface

The data used in this thesis was extracted from Glassdoor, an online third-party organizational review platform, with their permission. Part one of this thesis, which examined the impact of peripheral cues for Third-Party Online Organizational Reviews, was presented at, and published in the Academy of Management 2021 Conference proceedings.

Acknowledgements

Accomplishing this thesis would not be feasible without the input, expertise, and kind assistance from select individuals and entities. First, my deepest appreciation extends to my supervisor and mentor, Dr. Derek Chapman, who has been there every leg of this journey to offer his thoughtful critique and insights that paved the way for the conceptualization and execution of this project. I extend thanks to members of my supervisory committee, Dr. Joshua Bourdage, Dr. Adam Murry, and Dr. Ruth Pogacar for their thorough and highly valuable feedback. I would also like to thank reviewers from the Academy of Management conference who provided additional insights that helped to further improve and refine this project. I extend further appreciation to Michael McCarthy, from the Open Science Student Support Group (OSSSG), who willingly gave of his time to help me resolve an issue I experienced with one of the algorithms. OSSSG has also collectively provided the well needed resources that would facilitate making a reproducible script for other researchers to use. Further, I acknowledge and appreciate the feedback from graduate students in the Industrial/Organizational Psychology program on the coherency and labeling of topics. I would also like to thank Glassdoor for providing me with timely permission to scrape reviews from their platform, which helped to further facilitate ethical practices in data mining. And finally, I express deep thanks to my support system of family, friends, my partner, and Yahweh for cultivating a space that was conducive to executing this thesis from beginning to end.

Dedication

This thesis is dedicated to my family: my Godsister, Latoya Hall (laid to rest, June 2020), my Grandfather, Lloyd “Papa” Ingram (laid to rest, May 2021) and my Grandmothers, Rachel Alberta Harris (laid to rest, July 2011) and Daisy Ingram; my parents, Franklyn, and Jennifer Morgan; my siblings, Craig and Jamilia Morgan; my Godmother, Leleth Stewart, best friend, Kimberly Leckie and partner, George Imevbore.

Table of Contents

Abstract	i
Preface	ii
Acknowledgements	iii
Dedication	iv
Table of Contents	v
List of Tables	vii
List of Figures	viii
List of Abbreviations	ix
CHAPTER 1 Introduction: What makes a Helpful Online Organizational Review?	1
Review Helpfulness	4
Review Helpfulness for Consumer Reviews	5
Review Helpfulness for Employee Reviews	6
Elaboration Likelihood Model and Subjective Value Maximization	9
Elaboration Likelihood Model	9
Subjective Value Maximization	11
Review Sentiment	12
Review Sentiment and Helpfulness	12
Review Sentiment and Topic Proportions	14
Corporate Image and Review Consensus	15
Corporate Image and Helpfulness	15
Sentiment Deviation from Consensus and Helpfulness	17
Review Source	19
Review Content	22
TPOR Topics	22
TPOR Topics and Helpfulness Ratings	23
TPOR Components on Helpfulness Ratings	24
Topic Modeling	26
CHAPTER 2: Method	31
Dataset	31
Procedure	32

Analysis	34
Preliminary Analysis Part 1	34
Main Analysis Part 1	35
CHAPTER 3: Results	37
Part 1: Multiple and Moderation Regression Results	37
Part 2: Topic Modelling Results	43
Overview of Reviews	43
Pre-Processing	44
Search for k (number of topics)	46
Topic Modelling	50
Sentiment Analysis	58
Predictions	59
CHAPTER 4: Discussion.....	70
Overview of Main Findings and Theoretical Implications	70
Part 1 – Peripheral Cues	70
Part 2 – Topic Modelling.....	76
Practical Implications	83
Limitations and Future Research.....	86
Conclusion	89
References	90
Appendix A: Matching Labels to Terms Exercise.....	102
Appendix B: Interaction Plots with Topic Proportions and Sentiment.....	103
Appendix B1. Work Shifts/Hours	103
Appendix B2: Vacation and Overtime	104
Appendix B3: Company Culture/Climate	105

List of Tables

Table 1. Study's variables with their corresponding measurements.....	36
Table 2. Descriptive Statistics and Correlations for Part 1 Study Variables	38
Table 3. Results from moderation analyses of the relationship between sentiment and helpfulness ratings using overall ratings, deviation, and employee status.....	42
Table 4. OLS Multiple Regression Coefficients for Part 1 Study Predictors on Helpfulness Ratings	43
Table 5. Comparing the top words that emerged from preliminary analyses across the entire sample of reviews, and holdout samples.....	46
Table 6. Topics with their labels, terms, example comments and proportions.....	51
Table 7. Correlational matrix of helpfulness ratings, sentiment, and the topics' proportions	56
Table 8. Correlation matrix for the sentiment dictionaries and sentiment.....	60
Table 9. Correlational matrix of helpfulness ratings, sentiment, and the topics' proportions	64
Table 10. OLS Hierarchical Regression Results for Part Two Study Variables	65
Table 11. Part two summary of omnibus results for simultaneous regression models.....	66
Table 12. Correlation matrix for variables used in post-hoc analyses	68
Table 13. Summary of omnibus results for hierarchical regression models for the post-hoc analyses	68
Table 14. OLS Hierarchical Regression Results for post-hoc analyses.....	68

List of Figures

Figure 1. Relationship between individual and helpfulness ratings moderated by overall company ratings.....	39
Figure 2. Relationship between individual and helpfulness ratings moderated by sentiment deviation.....	40
Figure 3. Relationship between individual and helpfulness ratings moderated by employee status	41
Figure 4. Word cloud of most prevalent words across reviews, controlled for frequency	47
Figure 5. Word cloud of most prevalent words used by current and former employees	48
Figure 6. Word Cluster plot of words that co-occur with each other	49
Figure 7. Effect of employee status on topic prevalence	58
Figure 8. Relationship between Work Life Balance’s topic proportions moderated by sentiment	62
Figure 9. Relationship between Growth Opportunities’ topic proportions moderated by sentiment	63
Figure 10. Relationship between Leadership Quality’s topic proportions moderated by sentiment	64
Figure 11. Relationship between maximum proportion and helpfulness ratings at different levels of word count	69

List of Abbreviations

Abbreviation	Definition
ELM	Elaboration Likelihood Model
e-WOM	Electronic Word-of-Mouth
OCR	Online Customer Review
TPOR	Third-Party Online Organizational Review
TPPs	Third-Party Platforms (for employee reviews)
SVM	Subjective Value Maximization
WOM	Word-of-Mouth

CHAPTER 1 Introduction: What makes a Helpful Online Organizational Review?

Researchers have long been interested in attitudes toward organizations as well as identifying the mechanisms through which this occurs (Chapman et.al.2005; Rynes & Barber, 1990; Sung & Lee, 2015). Most researchers in that domain, focus on how employers can actively manage their company recruiting image through efforts like branding or impression management tactics (Cable & Judge, 2003; Yu, 2019). However, corporate recruiting images are also susceptible to influence beyond the control of the organization. Hearsay, according to Highhouse et.al., (1999), has been identified as one of the major contributors to company images, influencing attitudes such as applicant attraction. While companies have attempted to control their images through job advertisements and their recruiting websites, hearsay (or word-of-mouth (WOM)) has been given a vast electronic audience through online third-party platforms (TPPs) for employer reviews; thus, partially stripping organizations' recruiting agency (McFarland & Ployhart, 2015). TPPs have gained wide acceptance in the marketplace, with some major players such as Glassdoor and Indeed receiving between 60-250 million unique visits per month (Glassdoor, 2020; Indeed, 2020). It is therefore important to consider and examine the properties of TPPs alongside their potentially influential capacities (Van Hoya & Lievens, 2009).

TPPs have similarly gained wide traction through sites such as those that invite reviews on products, services, and experiences (for example, restaurant or hotel visits) (Filieri, 2015; Lee & Cranage, 2014). These reviews have fundamentally shifted the discourse of organizational images, as they have significantly contributed to how persons perceive a company through their product offerings (Lee & Cranage, 2014). Some platforms, such as Yelp, receive 178 million unique monthly visits (Bassig, 2018). Therefore, online reviews are inherently influential as they broadcast people's experiences (Van Hoya & Lievens, 2009) and provide insight on whether

interested individuals should pursue experiences with the company or product in question (Lee & Cranage, 2014). Because of this traction, the marketing literature has paid more attention to organic online customer reviews (OCRs) and what contributes to their influential nature (Filieri, 2015, Lee & Cranage, 2014; Ludwig et.al., 2013).

Though employee review TPPs have also dramatically altered recruitment processes, their impact remains largely unaddressed in the recruiting literature (Pernkopf & Wolfgang, 2020). Considering that TPPs symbolize a transparent source of organizational information for job seekers, there are certain implications for applicant attraction (Chapman & Gödöllei, 2017; Carpentier & Van Hoyer, 2020; Stockman et.al., 2020). For instance, because strong candidates have been found to be more influenced by recruiting information (Chapman & Webster, 2006), negative third-party online organizational reviews (TPORs) are a potentially stronger deterrent against highly skilled job candidates than weaker applicants. Moreover, it is the newer, lesser-known organizations as well as low to average rated companies with negative TPORs that are especially susceptible to losing traffic from more capable applicants (Stockman et.al., 2020).

However, there are certain presenting variables that may be able to buffer the effects of these negative reviews (Lee & Cranage 2014; Filieri, 2015; Van Hoyer & Lievens, 2007). A combination of peripheral and central cues can serve to either challenge or further reinforce the review's credibility and level of persuasiveness (Carpentier & Van Hoyer, 2020; Stockman et.al., 2020). Depending on the impact of these components on the review's credibility, candidates may be persuaded to either adopt a predominantly positive, negative, or neutral impression of the organization (Carpentier & Van Hoyer, 2020; Stockman et.al., 2020). This information adoption is potentially indicative through "helpfulness" ratings on the reviews, which describes the extent

to which viewers consider information in reviews useful for information processing and decision-making (Filieri, 2015).

While a few studies in the recruiting domain have explored the impact of different types of review sentiments on certain recruiting outcomes (Carpentier & Van Hove, 2020; Stockman et.al., 2020), to date, there has not been an empirical assessment of the employee reviews that organically emerge on TPPs, and their association with review helpfulness (Pernkopf & Wolfgang, 2020). Furthermore, the impact of the different components of organic TPORs on helpfulness ratings have not been meticulously explored. Employee review TPPs arguably act as a mechanism through which employee voice is amplified. Employee voice describes employees' engagement in informal communication about concerns regarding work-related issues to those who can initiate the necessary changes (Morrison, 2014). Further, employee voice can pose serious implications for the company's recruiting image, as sharing concerns online have been shown to affect applicant attraction (Stockman et.al, 2020). It is therefore important to begin the process of further assessing these elements and understanding the scope of their effects.

More specifically, this study serves to accomplish the following objectives: i) Examine the roles and impact of TPORs' sentiment, organizational ratings, consensus, and employee status, on helpfulness ratings; ii) evaluate the topics discussed in TPORs and their clusters, and iii) assess the extent to which the extracted topics impact helpfulness ratings. This mixed-method research inquiry is among a few in the recruiting arena that assesses thousands of organic employer reviews across hundreds of companies. Additionally, organic TPORs and their association with helpfulness ratings have not previously been examined. Accordingly, this research will offer an in-depth review of what contributes to an influential job TPOR. To accomplish this goal, we will examine at least 21,000 TPORs across 450 companies' profiles.

This will be achieved through utilizing automated web-scraping of online reviews, which are then submitted to automated text analysis tools, designed to efficiently extract, and assess an extensive amount of data.

Considering that reviews are inherently persuasive (Van Hove & Lievens, 2009), there may be certain reviews that lend themselves to be appraised as more helpful on these platforms than others. Therefore, we first assess the conceptualization of review helpfulness, the construct's relationship to information adoption, and the factors that tend to influence helpfulness. Then we examine the theoretical underpinnings of the Elaboration Likelihood Model and Subjective Value Maximization in the receptivity to, and persuasion by online reviews. We then evaluate how review sentiments influence helpfulness ratings, and how organizational ratings, sentiment deviation and employee status moderate the strength of that relationship. Finally, we evaluate the topics discussed in employee reviews and their relation to helpfulness ratings, before giving an overview of topic modeling.

Review Helpfulness

Online reviews, both in the recruiting and marketing domains, are typically accompanied with helpfulness ratings. Helpfulness ratings are votes from readers that indicate that they regard the communicated information as valuable and/or useful in effectively evaluating the described products, services, or experiences (Zhu et.al., 2014) they intend to procure. Because these ratings are considered an indication of collective consensus (Deng et.al., 2020; Risselada et.al., 2018), and by extension, review credibility (Cheung et.al., 2009), consumers are likely to prioritize processing reviews with higher votes. As a result, helpfulness ratings arguably act as heuristics for facilitating decision-making and influencing subsequent sales as they mitigate information overload on review platforms (Cao et.al., 2011; Zhu et.al., 2014). They also provide a marker for

how influential a given review is to prospective consumers/employees (Filieri, 2015; Zhu et.al., 2014). While helpfulness ratings can be assessed as a determinant of important outcomes, the current study specifically evaluates their antecedents, in light of the paucity of research in this area. Accordingly, parallel efforts in the marketing literature (Risselada et.al., 2018; Zhu et.al., 2014), are used to inform the conceptualization and potential antecedents of helpfulness in the recruiting domain.

Review Helpfulness for Consumer Reviews

This study focuses on company reviews and their potential impact on review helpfulness, however there are similar efforts in the marketing literature that provide insight into what information readers consider to be helpful. Online customer reviews (OCRs) for products and services have been thoroughly evaluated in the marketing domain and several studies have identified some of the elements that contribute to helpful OCRs (Filieri, 2015; Filieri et.al., 2018; Filieri et.al., 2019; Lee & Cranage, 2014; Li et.al., 2019; Sung & Lee, 2014). According to Filieri (2015), customer reviews are considered helpful if the information shared is relevant (i.e., applicable in evaluating the product/service/experience) and diagnostic (i.e., assist the consumer in familiarizing themselves with a product which would allow them to be able to remotely assess its quality and performance). These factors collectively contribute to how helpful the review is in facilitating effective information processing and decision-making (Zhu et.al., 2014).

While helpfulness ratings can be indicative of heuristic processing (i.e., leveraging simple cues to shortcut information processing; Filieri, 2015; Petty & Cacioppo, 1986), and not exactly helpful information, it is likely that reviews that are relevant and diagnostic will be considered more useful to consumers. Once information is regarded as helpful in the evaluation of the communicated message, this presents important implications for information adoption (Filieri,

2015). Information adoption describes the process of integrating the review's information into one's mental model and using that information for evaluative purposes. If the review contains a recommendation, then this process would also extend to accepting and adhering to these recommendations. Because regarding information as helpful (i.e., through relevance and diagnosticity) is a critical component of information adoption (Filieri, 2015; Sussman & Siegel, 2003), we use helpfulness ratings to infer this outcome.

There are a number of factors that influence a message's level of diagnosticity that serve to enhance or detract from OCRs' persuasive quality (Filieri, 2015; Filieri et.al., 2018; Filieri et.al., 2019; Li et.al., 2019; Sung & Lee, 2014). Specifically, characteristics such as the review's sentiment, information quality, relevance, length, readability, overall product rankings, company's responses to reviews and type of online comment (whether one-sided, two-sided and/or refutational) are typically examined as either moderators or direct influencers of helpfulness (Filieri, 2015; Filieri et.al., 2018; Filieri et.al., 2019; Li et.al., 2019; Sung & Lee, 2014). While these studies illustrate that, consumers prefer information that facilitates the message's diagnosticity, thus reducing ambiguity in their evaluations, the patterns may reflect differently on employer TPPs.

Review Helpfulness for Employee Reviews

Filieri (2015) argued that if the information presented facilitates the evaluation of where the product, service or experience stands in terms of quality and performance (i.e., information diagnosticity), then that influences how helpful the review is and by extension, information adoption. This may be applied to recruiting where a helpful review would be regarded as providing relevant and diagnostic information that allows other job seekers to effectively evaluate the nature of the organization. This would then offer insight on the extent to which the

job advertised is of a certain standard; thus, by indicating that a review is helpful, it means that the reader is able to make that determination through examining the online accounts (Filieri, 2015; Filieri et.al., 2018; Filieri et.al., 2019). Therefore, a helpful review likely facilitates acquaintance with the organization through its influence on information relevance, diagnosticity, and adoption (Filieri, 2015; Filieri et.al., 2018; Filieri et.al., 2019; Lee & Cranage, 2014).

Oppositional arguments for regarding a review as helpful are that disgruntled former employees are motivated to increase the consensus of negative organizational reviews. Alternatively, positive reviews could be marked helpful through initiatives from the organization to enhance their online image. Therefore, helpfulness ratings are potentially indicative of engagements from sources external to prospective applicants, introducing statistical noise into what these ratings reflect. While these arguments may be true, Glassdoor (2021) found through their US Site survey that 83% of their users are active jobseekers, and 89% highly regard employee accounts as critical in retrieving important information about a company. These findings illustrate that most reviews are likely perused by prospective applicants – whether these reviews are authentically created by actual current/former employees, is not under examination in this study.

While the marketing literature provides important insight into what consumers assess in their processing of reviews, different components may be responsible for certain reviews being more helpful than others on employee review TPPs. Persons arguably increase their evaluation of job reviews when compared to product reviews because the consequences are weighed heavier for prospective employment (Kruglanski, 1990; Pernkopf & Wolfgang, 2020). Moreover, organizational reviews typically feature accounts of current or former employees' experiences within a company, and experience reviews are associated with a heightened evaluation of

peripheral cues such as source credibility (Zhu et.al., 2014). Conversely, Wang et.al. (2019) show in their meta-analytic findings of helpfulness antecedents for OCRs, that source credibility was one of the least influential components, and that central cues such as information quality received greater evaluation. Thus, it can be inferred from these findings that source credibility may be valued differently between employees and consumers, and this provides impetus for deriving that review components may be weighed differently between employees and consumers.

Further, considering that reviewers are generally given the option of disclosing their identities on product review platforms, the act of doing so implicitly drives a level of trust among readers of that OCR (Filieri et.al., 2019). Additionally, because reviewers on these platforms often have profiles that depict their track record of reviews, consumers are able to determine their credibility and expertise in that domain (Filieri et.al., 2019). Moreover, the location of the reviewer, whether local or foreign, and whether they include photographs in their accounts are also deemed as helpful to potential consumers (Filieri et.al., 2018).

However, on TPPs for organizational reviews, reviewers do not have personal profiles nor are they able to disclose their identities. Therefore, the way in which viewers may assess for credibility and trustworthiness for the OCRs' reviewers would be different for employer TPPs. While we do not make direct comparisons between consumer and employee review platforms in the current study, we evaluate the impact of both peripheral (i.e., review sentiment, corporate image, sentiment deviation and message source) and central cues (i.e., topics discussed) for reviews, on helpfulness ratings. Previous research shows that when prospective consumers engage in complex processes to evaluate others' experiences, they evaluate both peripheral and central cues (Filieri & McLeary, 2013). Considering that employee and consumer review TPPs generally present different user layouts, it is important to examine how specific components are

evaluated in the recruiting domain. We further contextualize these arguments on the underlying mechanisms that likely increase the persuasiveness of, and receptivity to employee reviews, using the Elaboration Likelihood Model (ELM) and Subjective Value Maximization (SVM).

Elaboration Likelihood Model and Subjective Value Maximization

Elaboration Likelihood Model

The ELM is widely used to explain the mechanism through which a potential applicant is persuaded by a recruiting message (Gregory et.al., 2013; Jones et.al., 2006). The Elaboration (EL) continuum describes the level of processing and examination one engages in to dissect a communicated message and then determine the perceived value in adhering to it. The likelihood that a person will elaborate their processing of shared information depends on their levels of motivation and ability. Motivation refers to one's alacrity to thoroughly evaluate the reviews being conveyed whereas ability describes an individual's cognitive capacity or opportunity to painstakingly examine the information presented (Petty & Cacioppo, 1986). The ELM further proposes dual processes in elaboration, where the peripheral route emphasizes the assessment of cues that are extraneous to the core message, due to having low motivation and/or ability. Therefore, readers leveraging the peripheral route would shortcut information processing by relying on factors such as review consensus, message source and sentiment. In addition to the peripheral route, the central processing route involves the evaluation of informational cues at the message's core. Therefore, prospective applicants with high motivation and ability are likely to scrutinize the type and quality of arguments presented (Petty & Cacioppo, 1986).

Underscoring the arguments for what information readers both pay attention to, and are persuaded by, is the assumption that majority of the audience that evaluate reviews on employee review platforms are prospective applicants or active job seekers (Glassdoor, 2021). Further, the

prospect of evaluating potential places of work, or career paths is consequential for the individual as this decision likely presents more important implications, and longer lasting consequences, than those associated with choosing a product, restaurant, or hotel (Pernkopf & Wolfgang 2020). Considering that both the comprehensive evaluation of personal accounts, and careful integration of this information into decision-making, call on the engagement of complex cognitive processes (Filieri & McLeay, 2014; Zhu et.al., 2014), it is likely that majority of readers are highly motivated to process experience reviews (Zhu et.al., 2014) for organizations. Zhu et.al (2014) further argue that because experience accounts in reviews presume that the reader is not previously familiar with the place described, they consequently lack the necessary cognitive framework to more effectively evaluate the information shared. In other words, while they may possess the motivation to process the review, they lack the ability to successfully do so.

According to Petty and Cacioppo (1986), high motivation and low ability precede the use of peripheral cues. Thus, Zhu et.al. (2014) demonstrate that to determine if someone's experience would likely reflect that of the potential consumer, consumers are motivated to evaluate the message source to infer the level of similarity they have to the reviewer. Therefore, a lack in ability to effectively evaluate information, and the increased motivation to determine shared similarity with the message source, infer an increased assessment of peripheral cues on employee review platforms. However, majority of active job seekers on Glassdoor are professionals, where 45% have at least 6 years-worth of experience in their respective fields. While they may not have comprehensive knowledge of a particular organization, their previous engagement in their careers and current/former companies, likely provide a useful mental model for evaluating reviews. Further drawing on the ELM, Filieri and McLeay (2014) show that when consumers are highly involved in the processing of reviews, as they tend to be with

travel/experience review platforms, they engage in a thorough assessment of both peripheral and central cues. This is consistent with ELM arguments as having both high motivation and ability do not always necessitate discarding peripheral cues (Petty & Cacioppo, 1986). In fact, highly involved individuals are likely to elaborate their processing of peripheral cues more than those less involved (Petty & Wegener, 1999).

Subjective Value Maximization

While the ELM highlights the factors that job seekers are attentive to when processing reviews, the SVM provides further insight into what contributes to the persuasive nature of these components. Specifically, SVM describes the tendency for individuals to be persuaded by messages that they consider personally valuable and that presents implications for their social networks (Falk & Scholz, 2018). Therefore, when individuals process information, they are preoccupied with cues of social influence (i.e., the collective consensus in the opinions shared by the messenger and how it influences their presentation in social groups), as well as self-relevance (i.e., the extent to which the message affects personal goals, or ability to engage in self-enhancement).

Ultimately, SVM proposes that the personal relevance of information, and its implications for one's fulfilment of social needs supersede the analysis of the arguments' credibility and/or line of logic. Self- and social relevance arguably drive the processing of and persuasion by communicated information. Therefore, persons are motivated to avoid negative social sanctions (Falk & Scholz, 2018; Howe & Krosnick, 2017). This is consistent with the ELM, which outlines that when the audience has a personal stake in the message shared, there is a heightened elaboration of that information (Petty & Cacioppo, 1986; Petty & Wegener, 1999). And considering that active job seeking is a process that presents dual implications for one's

self-enhancement and social networks, it is expected that prospective applicants are motivated to process a diversity of cues that ultimately provide more information about the domains of self-relevance and social influence. Therefore, we propose that because active job seekers are invested in reading and assessing reviews for insight, and decision-making purposes, they will leverage both peripheral and central cues during these processes. Consequently, we examine review sentiment (i.e., individual ratings), corporate image (i.e., overall company ratings), sentiment deviation (i.e., the deviation of sentiment from overall consensus) and message source (i.e., employee status) as peripheral cues, and the type of topics (and their proportions) reviewers discuss as a central cue.

Review Sentiment

Review Sentiment and Helpfulness

Part of what contributes to the persuasive power of TPORs is the review's sentiment. The sentiment of a review is typically determined through the individual rating that it is associated with (Ghasemaghaei et.al., 2018). Each reviewer is usually asked to rate their overall experience or satisfaction with an organization on a scale of 1 (poor/negative) to 5 (excellent/positive) (also referred to as individual rating or review sentiment). Reviews with lower scores that indicate a predominantly negative experience, are clothed in the perception that the communicator is invested in the best interests of the audience; they also leave little room for ambiguity as they further facilitate information diagnosticity (Filieri et.al., 2019). In other words, readers are clear about the stance/opinions of the reviewer. Thus, potential candidates may appraise these reviews as intending to protect them from poor personal and social consequences, which subsequently align with their subjective values (Falk & Scholz, 2018).

While Pernkopf & Wolfgang (2020) found that their focus group participants were more skeptical of employee TPORs that were extremely positive or negative, their participants still acknowledged that TPPs were helpful because they presented unbiased reviews, especially those that were unpleasant. This is consistent with empirical studies that found that extremely negative reviews are not always helpful as their effects are often contingent on certain moderators (for example, size of company, review quality, length, source credibility; Cheung et.al., 2009; Evertz et.al., 2019; Filieri et.al., 2018; Filieri et.al., 2019; Stockman et.al., 2020; Wang et.al., 2018; Wu, 2013). Filieri et.al. (2019) further highlighted that experience reviews were more helpful to consumers when they were accompanied with moderate ratings. However, while reviews with mixed narratives (i.e., they are not predominated with one sentiment or opinion), may be perceived as credible information (Cheung et.al., 2009), they are more difficult to evaluate (Petty & Cacioppo, 1986), and are arguably not as impactful on helpfulness ratings (Cao et.al., 2011).

Instead, the marketing literature has consistently found that negative OCRs are generally associated with being more helpful (Cao et.al., 2011; Carpentier & Van Hoyer, 2020; Filieri, 2015; Filieri, 2018; Filieri, 2019; Kusumasondjaja et.al., 2012; Lo & Yao, 2019; Stockman et.al., 2020; Vaish, 2008; Wang et.al., 2018). This aligns with the inherent negative biases that humans are wired with that render them susceptible to the increased perception and influence of negative accounts (Vaish et.al., 2008). Negative information is more salient and easier to diagnose than those considered positive (Cao et.al., 2011; Carpentier & Van Hoyer, 2020; Filieri, 2015; Filieri, 2018; Filieri, 2019; Kusumasondjaja et.al., 2012; Lo & Yao, 2019; Stockman et.al., 2020; Vaish, 2008; Wang et.al., 2018). Specifically, Kusumasondjaja et.al. (2012) and Lo and Yao (2019) found that negative information is perceived as more credible, and as a result attract higher

helpfulness votes; but positive reviews elicit higher initial trust among readers (Kusumasondjaja et.al., 2012).

Therefore, we expect a similar pattern to be found on employee TPPs, where prospective applicants would be attentive to the sentiment of reviews, such that negative reviews would be considered more helpful. As mentioned earlier, invested applicants, such as those who are more skilled and competitive, are more sensitive to negative recruiting information (Chapman & Webster, 2006). These applicants are more scrutinous and selective of organizations they intend to apply to, and/or accept an offer from. This is indicative of their level of involvement in the processing of recruiting information about a company, especially given that their selection of an organization presents implications for their career growth. Similarly, active job seekers on third-party platforms are also invested in the process of reading reviews for organizations (Glassdoor, 2021), therefore they may further increase their evaluation of negative reviews, and subsequently consider that information to be more helpful than positive information.

Hypothesis 1: Negative reviews are perceived as more helpful than positive reviews.

Review Sentiment and Topic Proportions

We also anticipate that the sentiment of the review would interact with a topic's proportions. Topic proportion describes the extent to which a reviewer elaborates on a specific topic within the review. For instance, in a given review, there may be discussions about the organization's leadership and relationship with co-workers, however these topics could be discussed in varying amounts (for example, organizational leadership may dominate 70% of the review, and co-worker relations 30%). We currently focus on a specific topic within the entire review and examine the implications of the extent to which positive, negative, or neutral discussions about a specific topic, influence helpfulness ratings. Specifically, it is expected that

when negative information is highlighted and further elaborated in the discussion of a particular topic, it will receive higher traction than the discussion of negative reviews in smaller proportions for that topic (i.e., smaller proportion of negative sentiment indicates higher moderate to positive sentiment). Echoing the previous arguments made, a predominantly negative account of a topic would be considered easier to evaluate and diagnose (Filiari, 2015).

Hypothesis 2: Higher topic proportions associated with negative sentiments are perceived as more helpful than smaller topic proportions associated with negative sentiments.

Corporate Image and Review Consensus

Corporate Image and Helpfulness

Corporate image is defined as the aggregate of perceptions that external stakeholders hold about a company. It concerns the attitudes, emotions, and/or thoughts that consumers, and prospective applicants, associate with an organization (Cian & Cervai, 2014). Organizations have typically used a combination of assertive-defensive and indirect/direct tactics to construct and manage how their images are perceived by potential employees/consumers. Strategies such as organizational promotion, exemplification and prosocial behaviours have been used to curate images that serve to attract potential incumbents (Yu, 2019). Images that portray notions of integrity, corporate social responsibility (Tsai & Yang, 2010) and moral uprightness are what organizations generally strive to represent, considering that these virtues reflect a positive image which is inherently attractive (Yu, 2019). These approaches have also been used to restore and protect images that have been tainted through scandals and negative electronic WOM. Therefore, organizational image is critical in the process of attracting and retaining applicants (Tsai & Yang, 2010; Yu, 2019).

Allen, Otondo and Mahto (2007) outlined that one of the antecedents of organizational attraction is subjective considerations of the company's image and the extent to which an applicant perceives that there is congruence in values (fit). Therefore, potential applicants' attraction to a company are substantially influenced by brand portrayals especially if the organization presents a resilient image (Allen et.al, 2007; Chapman et.al., 2005). One mechanism through which these images are conveyed is through recruitment advertisements which influence organizational attraction; they are even more persuasive when the company's image aligns with the content in the advertisements (Baum et.al., 2016). Therefore, the brand of an organization can be the lens through which people sort and form attitudes toward a company (Allen et.al, 2007; Baum et.al, 2016; Cable & Turban, 2003; Chapman et.al., 2005; Stockman et.al., 2020).

On TPPs, a reflection of the brand image's resilience, and uprightness is expressed through an aggregate star rating of a company, which typically ranges from 1 (overall consensus among all reviewers is negative) to 5 (overall consensus among all reviewers is positive); potentially affecting the extent to which the effects of positive or negative reviews are either enhanced or dampened (Stockman et.al., 2020). A strong brand image communicates to customers and employees that the company possesses upstanding attributes that is constructed through a collective opinion, which has been found to be more impactful than a few negative opinions (Stockman et.al., 2020). This is consistent with the SVM, as indicators of social influence and consensus tend to evoke group conformity, where the overall agreement on the nature of the organization is inherently persuasive (Falk and Scholz, 2018). This is more critical if the overall company ratings are either positive or negative, as readers can derive what the collective opinion is. Otherwise, in cases where the overall organizational rating is more neutral,

prospective applicants may rely more on the information contained within individual reviews, especially reviews that are associated with negative sentiments.

Stockman et.al., (2020) illustrate that point, as they found that companies with higher mean ratings or stronger brands were found to be unchallenged when they receive negative reviews. Conversely, lower ranked organizations and those with weaker brand images were more susceptible to the influence of negative reviews, and experienced lowered applicant attraction. Therefore, this shows that prospective applicants may rely less on negative individual reviews, perhaps questioning their credibility, and instead prioritize consensus information that is indicated through the overall company ratings. Hence organizational ratings are expected to depict a buffering effect, where negative reviews for high rated companies will be considered less helpful, when compared to those low-rated.

Hypothesis 3: Negative reviews for high rated organizations are perceived as less helpful than those of low rated organizations.

Sentiment Deviation from Consensus and Helpfulness

Considering that the reviewers' consensus, which can be depicted through the overall rating of the company, may protect from the effects of negative reviews, this level of agreement may also impact potential incumbents in other ways. New attitudes formed while assessing TPORs can be reinforced and strengthened through signals of consensus (Howe & Krosnick, 2017). When reviews express similar sentiments, a consensus is created among current and former employees on the issues that a company primarily faces (Filiari, 2015; Könsgen et.al., 2018). And conversely, a review that deviates from the general sentiments contained within a review, is likely challenged on grounds of credibility, and considered less persuasive (Könsgen et.al., 2018).

Könsgen et.al. (2018) examined how the congruence of TPORs affect the viewers' assessment of the company's trustworthiness. They assessed the convergence of over 25,000 employee reviews on Kununu.de, a German TPP. When these were further evaluated in an experimental format, they found that divergence in accounts and sentiments decreased persons' trust in the company and by extension in organizational attraction. Therefore, divergent reviews express inconsistent imagery about the company in question, which is confusing for information processing and decision-making purposes. Furthermore, Filieri (2015) found that OCRs were more helpful (facilitated the ability to evaluate the product) when the sentiment of the product review was similar to that of others. This agreement among the sentiments expressed in reviews retains a persuasive quality (Filieri, 2015), and is further reinforced in the marketing literature where deviant information lacks credibility (Cheung et.al., 2009). In addition to Filieri (2015) determining that the congruence in sentiments is intrinsically persuasive, Pernkopf and Wolfgang (2020) demonstrated that mixed signals about an organization lowers employer attractiveness. This means that when applicants are exposed to divergent information about a company, they become ambivalent about the attributes and values that are upheld. And while this might enhance general informed decision-making, it disrupts the ability to determine an overall evaluation of the organization. These findings demonstrate that reliable information about a company enables prospective applicants to make decisions that align with the perceived consensus. We therefore anticipate that if a review deviates in sentiments from the overall consensus, reviewers would consider the information as confusing, inconsistent, untrustworthy, and consequently less helpful. Thus, it can be inferred that if the sentiments of the review align with those of other TPORs then the review may be perceived as more reliable and henceforth, more helpful.

Hypothesis 4: Reviews with similar sentiments to other reviews, will receive higher helpfulness ratings than reviews that deviate in sentiment.

While the consensus among reviews is an important predictor of helpfulness ratings, source credibility and the content of the message are also conventionally appraised as being crucial in persuasion. Therefore, these elements may have considerable implications for the extent to which reviews receive helpfulness ratings.

Review Source

TPPs, when compared to web-based employee testimonials, are inherently more persuasive, according to research from Van Hoyer and Lievens (2007). Web-based employee testimonials are orchestrated by organizations, and feature reviews from employees on the company's website. The information shared by these employees typically serve to bolster organizations' image, therefore are less likely to feature negative word-of-mouth (Van Hoyer & Lievens, 2007). When compared to employee review TPPs, employee web-based testimonials are less persuasive because TPPs present anonymous reviews from current and former employees; thus, there is no obligation for these reviewers to exaggerate the images of their places of work (Saini & Jawahar, 2019). TPORs' authors are therefore given a platform to share their experiences with little to no repercussions for doing so (Van Hoyer & Lievens, 2007). Furthermore, job candidates are generally motivated to seek accounts from these sources that have an intimate knowledge of the companies as they are perceived as providing more accurate and transparent information (Van Hoyer & Lievens, 2009).

This research substantiates arguments that prospective applicants evaluate message source and consider the implications of the message source's reviews, and credibility. Thus,

being more inherently influenced by TPPs further shows that persons are attentive to cues surrounding source credibility of TPORs (Petty & Cacioppo, 1986; Van Hove & Lievens, 2007; Van Hove & Lievens, 2009; Filieri et.al., 2019). Given this, those who are currently or previously affiliated with an organization are thought as having an intimate knowledge of that space. However, it is also possible that attentiveness to source credibility also extends to the reviewer who shares the message. This is similar to product/experience reviews, where consumers evaluate the reviewer's ranking, number of reviews shared and location (not exhaustive), to infer credibility (Filieri et.al., 2018; Filieri et.al, 2019). Similarly, prospective applicants may be attentive to the reviewer's credibility through cues such as their employee status. Though source credibility was found to be less influential than factors such as information quality and sentiment in the consumer literature (Wang et.al., 2019), it was considered an important element for evaluating experience reviews (Zhu et.al., 2014). Source credibility specifically presents an elevated importance in "experience" reviews, which are similar to employee accounts. Specifically, persons search for information that would assess the extent to which they are similar to the reviewer (such as checking the reviewer profile, and location), as someone's experiences are heavily influenced by their unique perceptions (Zhu et.al., 2014). Therefore, it is likely that employer TPPs will demonstrate a heightened evaluation of the message source's credibility, considering that these reviews center on employee experiences and accounts.

Pernkopf and Wolfgang (2020) found through focus groups that the anonymity of employee reviews inherently raises suspicion among job seekers regarding source credibility. This is due to the difficulty with assessing this facet of a review if one's identity is not disclosed. These concerns are further reinforced when participants in the focus group revealed that they

perceive negative reviews from former employees as less credible, because former incumbents may have experienced an unsatisfactory discharge, potentially biasing their experiences (Pernkopf & Wolfgang 2020). It may also be the case that job seekers assess a negative review from a current employee as indicative of the present work climate whereas negative feedback from former employees may be perceived as dated. Therefore, negative appraisals from a current employee in this context would perhaps be evaluated as more reliable and helpful information.

Conversely, Saini and Jawahar (2019) found support for the notion that current and former employees have differential experiences (Sainju et.al., 2021) that impact their recommendations of a company. Because former employees have lower stakes in the organization, and as a result, are less impacted by the fate of the company, they are likely more transparent in their accounts. On the other hand, current employees are likely to recommend their companies as a desirable place to work (Saini & Jawahar, 2019), which suggests that their reviews are more positively oriented. This strongly implies that if former employees' reviews are less positive, and readers are inherently biased towards processing and accepting negative information (Vaish, 2008), then they may determine former employees' reviews to be more helpful. Therefore, given the conflicting results from the qualitative and quantitative approaches, the direction of this association is not apparent. We therefore raise the question of how employee status impacts helpfulness ratings, and whether it interacts with review sentiment. Considering that employee status tends to be examined alongside the sentiment of their reviews (Pernkopf & Wolfgang, 2020; Saini & Jawahar, 2019), we consider both factors in their collective influence on helpfulness ratings.

Research Question 1: How will employee status impact helpfulness ratings (1a) and interact with review sentiments in its effects on helpfulness ratings (1b)?

Review Content

TPOR Topics

Cao et.al., (2011) proposed and illustrated, through their use of text mining techniques, that the semantic aspects of a customer review are most essential in attracting helpfulness ratings. Therefore, it is critical to assess the actual content of reviews as an important component of TPORs. Previous research employing automated text analytic techniques on organic organizational reviews, have illustrated how certain topics discussed on employee TPPs provide insight into job satisfaction factors, and their implications for turnover (Jung & Suh, 2019; Lee & Kang, 2017; Luo et.al, 2016; Sainju et.al., 2021; Stamolampros et.al., 2019) and corporate performance (Luo et.al., 2016). The topics that were extracted across these studies were very varied, and at times were low-level (therefore, not exactly encapsulating higher order themes) but generally concerned management (and at times discussed in conjunction with leadership or employee well-being), work life balance, pay and benefits, organizational culture, and career advancement (Jung & Suh, 2019; Lee & Kang, 2017; Luo et.al, 2016; Sainju et.al., 2021; Stamolampros et.al., 2019).

Further, Dabirian et.al (2017) examined 38,000 Glassdoor reviews for companies that were either ranked as best or worst places to work. They found that there were certain topics discussed more frequently from employees of top-ranked organizations than those of low ranked. For instance, former/current employees of high-ranked companies were most satisfied with how enjoyable the workplace is. However, the ability of the organization to substantially compensate their employees was under most scrutiny from those of the worst-ranked companies. Overall,

topics related to the social, interest, economic, application, work/life balance and developmental values of the company, were discussed. These studies illustrate the general themes highlighted on employer review platforms and provide insight into what topics we expect to be extracted. Considering that the text mining of employee reviews is a relatively novel pursuit in the recruiting literature, we aim to further explore this facet of employee TPPs, as there may be additional topics not captured in prior research. Considering that most of this research was conducted in the Computer Science domain, we hope to further conceptualize these themes against theoretical frameworks to improve our understanding of what concerns reviewers tend to cover, and readers are influenced by. We randomly selected companies based in the United States and Canada (therefore, may be differing from samples used in other studies), and raise the question of what topics are typically discussed on employee review platforms, through these lenses.

Research Question 2: What topics are typically discussed on employer review platforms?

TPOR Topics and Helpfulness Ratings

SVM also bolsters the argument that certain topics of contention/praise in TPORs may receive more traction and be more helpful than others, depending on the needs of the audience. Considering that persons are more convinced by and/or perceptive of information that is meaningful to them (Falk & Scholz, 2018), it is likely that certain topics discussed in TPORs will be considered more critical to assess than others. This is observed in Chapman et.al. (2005) meta-analysis on the factors that affect applicant attraction where work environment and organizational image were considered the most influential predictors of job-organization attraction, whereas other factors such as compensation/benefits were found to be less influential. Likewise, Turban (2001) illustrated that a firm's attractiveness was largely influenced by

organizational attributes, particularly those concerning Challenging Work and Company Image. However, Compensation and Job security was not found to be influential as an organizational attribute, which further implies that prospective incumbents would be especially keen to assess how reviewers address certain facets of a company, more so than others. And as a result, if reviewers highlight certain topics of interest in greater proportion than others, readers may find those to be more helpful in their decision-making processes. Thus, given the many topics that are presumably discussed on these platforms and that previous research has not examined their effects on helpfulness, we raise the question if there are certain issues that may resonate with potential applicants more than others, as indexed by the number of helpfulness votes associated with the review.

Research Question 3: What topics are typically regarded as more helpful in online employee reviews?

TPOR Components on Helpfulness Ratings

As illustrated in previous studies in the marketing domain, there are certain components of an online review that may be more influential than others. In a meta-analysis on factors that contribute to a helpful consumer review, Wang et.al., (2019) found that the quality of information shared (as measured through review depth) was most critical in facilitating decision making, followed by the sentiment and age of the review. Filieri's (2015) study was also assessed in this meta-analysis, where source credibility was found to be least helpful to consumers, which raises the question of how these factors are perceived in an organizational review context. Like the studies conducted in the consumer literature, we also anticipate that there may be certain components of an employee review that potential job seekers would regard as more influential than others.

Though information quality was excluded from our analyses, the patterns generally found for product review platforms may not reflect for employee TPORs. Considering Pernkopf and Wolfgang's (2020) arguments that persons on company review sites are more evaluative of information posted there, we can anticipate graver implications for choosing a place of work than a product, potentially driving the motivation to prioritize examining certain components over others. Therefore, there is an expectation that TPOR factors will be differentially influential on helpfulness ratings, as reflected in the consumer literature (Wang et.al., 2019). We previously argued that prospective applicants will pay attention to both peripheral and central cues in their examination of reviews, however it is difficult to determine where exactly those differences will reside, in terms of what active job seekers are most attentive to. Considering the possibility that employee reviews may receive different elaborations in processing than consumer accounts, we instead raise the question of how TPOR components will compare in terms of their influence on helpfulness ratings.

Research Question 4: What TPOR components will receive higher helpfulness ratings in online employee reviews?

This large-scale assessment of employee TPORs serves to better understand the extent to which different components of a review influence someone's appraisal of a company. This involved examining the effects of the sentiment, overall rating of the organization, the extent to which reviews' sentiments are different from other reviews of the company (i.e., deviation), the type of reviewer (former or current employee) and topics discussed, on helpfulness ratings. An automated text analysis using Latent Dirichlet Allocation (LDA), alongside Structural Topic Modeling (STM) were used to determine the themes/topics discussed in the TPORs.

Topic Modeling

Topic Modeling is specifically used to retrieve and classify the words (similar to codes in thematic analysis) used in chunks of text, into topics (commensurate to themes in thematic analysis, or factors in more traditional empirical analyses) (Banks et.al; Baumer et.al., 2016; Buenaño-Fernandez et.al., 2020; Doldor et.al, 2019; Li et.al., 2019). These topics can then be used to predict certain outcomes of interest (Banks et.al., 2018), such as helpfulness ratings in the current study. LDA is the most prevalent topic modeling approach that is unsupervised and fully automated in its analysis of text (Banks et.al., 2018; Denny & Spirling, 2018). Unlike the supervised methods of text analysis, the LDA does not allow the researcher to designate the input and output processes of the algorithm; instead, its results are refined through the analysts' feedback/expertise (Banks et.al., 2018; Denny & Spirling, 2018). For instance, the researcher may decide on the number of topics that best represent the dataset, but they would not be able to dictate how words/documents are categorized into topics.

The LDA utilizes a “bag-of-words” approach, which aligns with the underlying philosophy that the grammatical structure of words is not necessary to derive *what* is discussed in the text (Banks et.al., 2018; Buenaño-Fernandez et.al., 2020; Denny & Spirling, 2018; Doldor et.al., 2019; Liang, 2019). As a result, the topic modeling process reduces the documents to terms and removes stop words/qualifiers that do not substantially contribute to the meaning of the text (e.g., articles/stop words such as “the”, “a”), thus disrupting the grammatical structure. Words are stemmed and/or lemmatized (i.e., they have been reduced to their “base” or a simpler form (Banks et.al; Buenaño-Fernandez et.al., 2020; Denny & Spirling, 2018; McKenny et.al.,

2018). For example, “*cultur*” could be interpreted as “culture” and/or “cultural” and/or “cultured”) to mitigate analytical redundancies, thus improving the quality of topic extraction. Specifically, this avoids the inundation of the dataset with several words that are essentially related to each other; if they are not stemmed, the same root words may potentially be assigned to different categories, which complicates the interpretation of individual topics (as they may conceptually overlap with others) (Banks, 2018; Buenaño-Fernandez et.al., 2020). After documents are reduced to words, each word is assigned probabilities for belonging to each topic, and clusters of words are identified through their probabilities and co-occurrences with each other. It is this product that we would use to interpret the output of higher order topics or themes (Banks et.al., 2018; Baumer et.al., 2016; Buenaño-Fernandez et.al., 2020; Doldor et.al, 2019; Li et.al., 2019).

The LDA approach can be understood as an integration of thematic analyses (i.e., where text is deconstructed into codes and re-integrated into themes), and machine learning, (i.e., affording the extensive analyses of copious amounts of data) (Banks et.al., 2018; Baumer et.al., 2016; Doldor et.al., 2019). A combination of these approaches proposes a retainment of the integrity of thematic analyses that can be extended to benefit the assessment of thousands of datapoints. This is otherwise a difficult feat to achieve with traditional (i.e., manual) methods for text analyses (Baumer et.al., 2016; Doldor et.al., 2019). Like Exploratory Factor Analysis (EFA), each document in the dataset is assigned to the same themes where some documents have higher probabilities of belonging to that topic than others, based on the word usage in that topic (Banks et.al. 2018; Buenano et.al. 2020; Doldor et.al., 2019; Li et.al. 2019). The document-topic matrix indicates how much of the specific text is addressed by each topic (i.e., the proportion of each

document that applies to a specific topic), which can then be used to make important predictions (Banks et.al., 2018).

A significant drawback cited for topic modeling is that this technique does not capture the context in which certain words are used (Buenano et.al. 2020; Denny & Spirling, 2018; Doldor et.al., 2019; McKenny et.al., 2018). For example, a specific topic might include the word “pay” which may lead researchers to believe that reference is being made to compensation, in a dataset of employer reviews. However, an example of a text that uses the word “pay” may be, “this company will *pay* for their atrocities”. It is apparent here that “pay” in this context is not equivalent to compensation. To mitigate the misinterpretation of certain words within a topic, it is important to refer to the texts that included that term and note the way in which that word is predominantly used (Banks et.al., 2018). If an overwhelming majority of this word usage across texts is indeed indicative of compensation, then it is appropriate to default to the initial interpretation (i.e., compensation). These examples illustrate that topic modeling focuses on *what* is discussed but does not capture *how* it is discussed which aligns with the goal of the current study, as we are most interested in uncovering *what* reviewers highlight in their TPORs. We further reiterate this point by omitting all words from the analyses that serve to further contextualize the texts, such as indicators of sentiments/attitudes (for example: “terrible”, “horrible”, “good”, “great” etc.). The pros and cons of each review were also combined into one complete text to further dampen the sentiments contained in the accounts, as our goal is to examine the topics covered in reviews.

As previously mentioned, examining the texts in which terms were used will aid with determining the word’s context, which facilitates topic labeling. Word cluster plots, which are visuals of the words that generally co-occur with each other, also illustrate how certain terms are

typically discussed. Therefore, employing both approaches in examining the topics should prove effective with revealing the critical themes discussed by reviewers. Other contextual cues, such as the sentiments of reviews, are difficult to retrieve using topic modeling. Instead, natural language processing, such as that of sentiment analysis, is recommended for determining the sentiment of accounts shared (Banks et.al., 2018). Sentiment analysis was solely leveraged in this study to validate individual ratings as a proxy variable for review sentiment. Individual ratings of a review have been previously shown as indicators of sentiment in the consumer literature (Ghasemaghaei, 2018), but has not been verified in the context of employee reviews.

A significant benefit of the topic modeling approach is that it allows for the assessment of covariates – this is accomplished in a higher order structural topic model (STM) (Banks et.al., 2018). We use STM to retrieve the most stable topics from the dataset, as well as assess how topic prevalence is affected by the reviewer’s job status with the organization (i.e., what topics are most prevalent among current and former employees); this would otherwise be difficult to retrieve using traditional methods for text analysis (such as with manual thematic analysis). Baumer et.al. (2017) have attempted to cross-reference the results from the LDA with manual thematic analyses to examine the automated tool’s capacity to retrieve appropriate and coherent topics; these researchers were able to extract similar themes from both approaches, which suggests that the LDA is effective with capturing the most critical subject matters outlined in text.

The classification protocol, established through the topic modeling approach, will therefore serve to categorize TPORs in order to better understand the extent to which different components of a review influence someone’s appraisal of, and satisfaction with a company. The study’s objectives are achieved in two parts. Part one examines the peripheral components of the

review (review sentiment, corporate image, employee status and sentiment deviation). Part two evaluates the central components of the review – topics retrieved from the text analyses and how these topics influence helpfulness ratings and differ between employee statuses.

CHAPTER 2: Method

This study's primary aims were to examine TPOR components' impact on helpfulness ratings. This involved assessing reviews across companies' Glassdoor profiles and extracting the text, helpfulness, and review sentiment (i.e., individual ratings), as well as the employee status for each review and the respective companies' overall ratings for further analysis.

Dataset

We assessed companies' Glassdoor reviews, as Glassdoor is one of the most formidable employee review TPPs, engaging 67 million unique monthly visits (Glassdoor, 2020). Reviews that are recent and/or have higher helpfulness ratings, are likely to attract more helpfulness ratings (Zhu et.al., 2014; Filieri, 2015), which confounds the ability to assess the influential nature of TPOR components. This is especially concerning if Glassdoor's platform prioritizes showing reviews that are trending in helpfulness votes. However, Glassdoor does not organize reviews by helpfulness ratings, nor by their posted dates, which helps to mitigate these potential biases.

We collected 24,687 Glassdoor reviews where the entire dataset of extracted reviews was used for the topic modeling analysis in the second half of the study. However, we assessed 21,542 reviews across 450 companies' profiles, with approximately 50 reviews per organization, for part one of the study. Glassdoor reports that readers form an impression of the organization after reading approximately four reviews (Glassdoor 2020; Glassdoor 2021). Though this finding needs to be validated, it illustrates that scraping the last 50 reviews posted for companies likely reflects the information that prospective applicants would engage with. For the first part of the study, companies with no helpfulness ratings for any of their reviews were omitted from the

dataset, leaving a range of 51 to 41,000 reviews per company and helpfulness ratings that ranged from 0-18 (after omitting extreme outliers) per review. Forty-six percent of the Glassdoor reviews in this sample were provided by current employees where the average organizational rating was 3.53 ($SD=.42$). Reviews had a mean individual rating of 3.64 ($SD=1.23$) and were found to be deviating from their aggregate company score by 1.32 ($SD=.81$) points on average. Overview of the study variables and their measurement are shown in Table 1.

Procedure

First, we requested and received approval from Glassdoor to scrape reviews from up to 1000 companies from their platform. Wikipedia was then used to identify 3800 companies that are based in the United States of America and Canada. We then randomly selected 1000 companies from that sample and retrieved their respective URLs on Glassdoor. Octoparse, an open-source automated web-scraping tool, was used to collect data from the first 5 pages (i.e., 50 reviews) from companies' Glassdoor profiles, as conventional manual data collection would preclude the ability to incorporate both depth and breadth in the analyses. We programmed Octoparse to scrape the TPOR components aforementioned where 495 companies had at least 50 Glassdoor reviews (which was our criteria for this study). Thus, we extracted 24,687 reviews, which was less than the expected 24,750. This is because some companies were projected on Glassdoor as having a total of 50 reviews, however, they had less than 50 that were in English, which were not readily available to scrape on the site (non-English reviews are hidden unless the filter is manually changed by the visitor). Furthermore, other companies had no helpfulness ratings for their reviews. For these reasons, 43 companies were omitted from the dataset, leaving 22,638 reviews.

We then dummy coded employee status (0 assigned to current employees and 1 to former) and assigned identification numbers to each company in the dataset. Deviation scores were also computed as the squared difference between individual ratings and their respective aggregate company ratings (as it appears on Glassdoor). We verified if each company had 50 reviews and completed additional scraping to ensure that all organizations were equally represented in the dataset. The pros and cons for each review were combined into one document for each reviewer, to increase the breadth and focus of the analyses. The data underwent preliminary analyses, before being examined in moderation and simultaneous regression analyses.

The reviews were then examined in R (programming language for statistical analyses) to determine the type of topics discussed, and to extract the document-topic proportions for further analyses (tidyverse, quanteda and stm packages were used; R Core Team, 2013). The TPORs' readability (i.e., the ease of reading a chunk of text), word and character counts were first evaluated, as well as the word frequency plots. The TPORs were converted to lowercase, and were stemmed, and trimmed (once the reviews were converted to terms/tokens, terms that were either too infrequent or frequent were omitted from the dataset to prevent hindrance of interpretation).

The data was randomly split in half to evaluate if the top words recognized in the entire dataset were relatively the same across both holdout samples. After examining the consistency of words across the reviews, we then determined the number of topics appropriate for the entire dataset. This involved running several series of topic models to assess their semantic coherency (i.e., interpretability) and held-out likelihood (i.e., predictive capacity). These metrics were used to compare the quality of these topic models, to then determine which number of topics seem to

best represent the discussions observed in the dataset. Once there was consensus on the topic-model that was most appropriate for the data, the corpus was submitted to a STM analysis where employee status was entered as a covariate. The topics provided from the overall topic-model analyses were then interpreted using the words that had the highest probability of belonging to that topic (i.e., probability words) and words that were frequent and exclusive to that topic (i.e., FREX words). We also examined the representative text for the words associated with each topic to evaluate the context in which the words were used, to avoid misrepresenting the TPORs. Once the individual topics were labelled, we assessed their level of interrater agreement with three psychometrically trained graduate students in Industrial/Organizational Psychology. After refining the topics through feedback from the group of graduate students, we ran a cluster analysis through STM to assess the extent to which the topics observed are related to each other. These topics were then submitted to a hierarchical regression analysis, to assess their main effects, as well as the effects of their interaction with review sentiment, on helpfulness ratings. More details about the topic modeling analyses are discussed in the results section.

Analysis

Preliminary Analysis Part 1

The dataset was first examined for univariate outliers where a strict criterion was established that omitted outliers exceeding ± 2.58 SD from their respective means. This resulted in an additional 1095 reviews being omitted, leaving 21,542 in the dataset. The data was then assessed for normality in distribution, linearity, homoscedasticity, and normality of errors. Normality of distribution was determined through non-significant Kolmogorov-Smirnov tests and linearity was observed by creating scatterplots between each TPOR component and helpfulness ratings. We also examined homoscedasticity by constructing a scatterplot between

the Z scores for both the predictors and the residuals. Finally, the normality of errors for each predictor was plotted with a histogram. Zero order correlations between the TPOR components (helpfulness ratings, review sentiment, corporate image, deviation scores and employee status) were then computed.

Main Analysis Part 1

After bivariate relationships were computed, we examined overall company rating, deviation scores and employee status as moderators of the relationship between review sentiment (as indicated by individual ratings) and review helpfulness (as demonstrated through helpfulness ratings) using PROCESS macro developed by Hayes (2012). Further, we evaluated these TPOR components in a simultaneous OLS regression to glean how influential they are on helpfulness ratings. A simultaneous approach was chosen, as the exact influence of these TPOR components on helpfulness ratings is not apparent in the recruiting context. Otherwise, if it was previously determined that a specific component was the most influential among employee reviews, then we would leverage a hierarchical regression. A hierarchical regression was employed in part two of this study to examine if topic proportions, and their subsequent interaction with review sentiment, had incremental variance beyond that of the strongest predictor of helpfulness ratings identified in part one of this study.

Table 1

Study's variables with their corresponding measurements.

Study Variables	Measurement
Review Helpfulness	Helpfulness ratings for each review
Review Sentiment	Individual ratings for each review
Corporate Image	Aggregate of the individual ratings for each review
Sentiment Deviation	Computed as the squared difference between Individual rating and Overall Company Rating
Employee Status	Employee status for each review
Review Topics	LDA to extract the topics discussed with their document-topic and document-term proportions.

CHAPTER 3: Results

Part 1: Multiple and Moderation Regression Results

The primary aim for part one of this study was to examine how peripheral TPOR components influence helpfulness ratings and moderate the relationship between review sentiment and review helpfulness. Hypotheses 1, 3, 4 and Research Questions 1 and 4 are addressed in this section, where the remaining hypotheses and research questions are assessed in Part 2. Preliminary analyses revealed that the distribution of the data and errors for each variable did not approach normality – helpfulness ratings were positively skewed, and review sentiments were negatively skewed. The predictors instead demonstrated levels of linearity with helpfulness ratings. There were also levels of homoscedasticity shown for review sentiment, which were not as apparent with the other predictors. Thus, we caution that results from this study be interpreted within these parameters.

Hypothesis 1, which proposed that negative sentiments (as indexed by lower individual ratings) would attract higher helpfulness ratings, was supported with a moderate negative association ($r = -.327, p < .001$). The remaining correlations between the TPOR components, and their descriptive statistics can be found in Table 2. We then examined overall company ratings, deviation, and employee status as moderators of the relationship between review sentiment and helpfulness ratings (results shown in Table 3). Hypothesis 3 predicted that company's aggregate ratings (i.e., the overall company rating on Glassdoor) would operate as a buffer against the impact of negative reviews on helpfulness ratings. We found that this model significantly predicted and accounted for 11% of the variance in helpfulness ratings. As shown in Figure 1,

negative reviews were least impactful on helpfulness ratings when overall company ratings were highest ($b = -.34$, $SE = .017$, $p < .001$), compared to when moderate ($b = -.45$, $SE = .01$, $p < .001$) and low ($b = -.53$, $SE = .01$, $p < .001$); thus, illustrating a partial buffering effect.

Table 2

Descriptive Statistics and Correlations for Part 1 Study Variables (N=19,934)

Variable	Mean	SD	1	2	3	4
1. Review Helpfulness	0.62	1.78	-			
2. Review Sentiment	3.65	1.22	-.32*	-		
3. Corporate Image	3.54	0.42	-.14*	.36*	-	
4. Sentiment Deviation	1.31	1.44	.24*	-.29*	-.25*	-
5. Employee Status ^a	0.51	0.50	.04*	-.21*	.12*	-.06*

Note. * indicates that correlations were statistically significant at $p < .001$. Listwise deletion was used. ^a0=current employee, 1=former employee.

Hypothesis 4 proposed that reviews' ratings that deviated more from their respective overall company rating would attract less helpfulness votes. The overarching bivariate association between deviation scores and helpfulness ratings was positive and significant ($r = .251$, $p < .001$), suggesting that reviews that deviated more from consensus ratings were regarded as more helpful. We further probed deviation as a moderator of the relationship between individual and helpfulness ratings, which presented a significant omnibus model ($p < .001$, $R^2 = .13$). Furthermore, a small but significant interaction between review sentiment and deviation scores was found ($p < .001$, $R^2_{\text{change}} = .0005$). Figure 2 illustrates that the strongest negative influence of review sentiment on helpfulness scores was apparent when deviation scores were high ($b = -.419$, $SE = .01$, $p < .001$), compared to when low ($b = -.356$, $SE = .02$, $p < .001$) and moderate ($b = -.372$, $SE = .01$, $p < .001$). This indicates that negative sentiments attract more

helpfulness ratings when they deviate from consensus (i.e., overall company ratings), therefore further opposing the hypothesis. Therefore Hypothesis 4 was not supported.

Figure 1

Relationship between review sentiment and helpfulness ratings moderated by corporate image.

Research Questions 1a and 1b raised the questions of how employee status would be associated with helpfulness ratings and interact with review sentiment on helpfulness votes. After dummy coding employee status (0 = current employee, 1 = former employee), the main effect revealed that former employees' reviews were generally considered more helpful than reviews provided by current employees. Evaluating employee status as a moderator of the relationship between review sentiment and helpfulness revealed a significant model ($p < .001$, $R^2 = .11$), with an interaction between review sentiment and employee status ($p < .001$, $R^2_{\text{change}} = .003$). These results further demonstrated that the negative effect of review sentiment on helpfulness

ratings was more pronounced when reviews were provided by former employees ($b=-.556$, $p<.001$) than current employees ($b=-.380$, $p<.001$), as shown in Figure 3. In other words, reviews with negative ratings were especially influential from former employees. Figure 3 also illustrates that as reviews increase in their positive sentiments, they are considered more helpful when shared by current employees.

Figure 2

Relationship between review sentiment and helpfulness ratings moderated by sentiment deviation.

Finally, Research Question 4 raised the question of which peripheral component of a review would be more influential on helpfulness ratings. Considering the exploratory nature of this question, we ran a simultaneous regression that included review sentiment, employee status,

overall company ratings and deviation scores as predictors in this model, which was significant ($p < .001$, $R^2 = .13$). As seen in Table 4, we found that all components except for overall company scores ($\beta = -.007$, $p = .299$) were significant, and that review sentiment exhibited the strongest influence ($\beta = -.276$, $p < .001$), followed by deviation scores ($\beta = .163$, $p < .001$) and employee status ($\beta = -.007$, $p < .001$).

Figure 3

Relationship between review sentiment and helpfulness ratings moderated by employee status.

Table 3

Results from moderation analyses of the relationship between review sentiment and helpfulness ratings using overall ratings, deviation, and employee status.

Model	<i>F</i>	Predictor	Coefficient	<i>p</i>	LLCI	ULCI	<i>SE</i>	<i>R</i>	<i>R</i> ²	ΔR^2
1	889.98	Constant	5.15	<.001	4.56	5.73	.30	.33	.110	
		Review	-1.12	<.001	-1.27	-0.96	.08			
		Sentiment								
		Corporate Image	-0.82	<.001	-1.00	-0.65	.09			
	50.14	Interaction	0.19	<.001	0.14	0.24	.02			.003
2	1103.39	Constant	1.66	<.001	1.53	1.80	.07	.37	.130	
		Review	-0.35	<.001	-0.39	-0.32	.02			
		Sentiment								
		Deviation	0.28	<.001	0.24	0.32	.02			
	13.22	Interaction	-0.02	<.001	-0.03	-0.01	.01			.001
3	788.19	Constant	2.57	<.001	2.48	2.67	.05	.33	.106	
		Review	-0.56	<.001	-0.58	-0.53	.01			
		Sentiment								
		Employee Status ^a	-0.53	<.001	-0.68	-0.38	.08			
	76.67	Interaction	0.18	<.001	0.14	0.22	.02			.003

Note: Model 1 *N*: 21536; Model 2 *N*: 19934; Model 3 *N*: 21536. ^a0=current employee, 1=former employee.

Table 4*OLS Multiple Regression Coefficients for Part 1 Study Predictors on Helpfulness Ratings**(N=19,934)*

Variable	<i>b</i>	β	<i>t</i>	<i>SE</i>	<i>p</i>	VIF
Constant	1.88					
Review Sentiment	-.40*	-.28	37.29	.01	<.001	1.25
Overall Ratings	-.03	-.01	-1.04	.03	.299	1.18
Deviation Scores	.20*	.16	23.18	.01	<.001	1.12
Employee Status ^a	.12*	.03	4.74	.02	<.001	1.05
<i>R</i>	.36					
Adjusted <i>R</i> ²	.13					
<i>F</i>	726.30*					

Note. * Indicates that regressions were statistically significant at $p < .001$. Listwise deletion was used. ^a0=current employee, 1=former employee.

Part 2: Topic Modelling Results

Overview of Reviews

There were 24,687 reviews extracted from Glassdoor, upon permission provided from the employer review platform. These reviews represented a wide range of sentiments (i.e., positive, negative, and neutral attitudes) as well as were characterized by varying lengths, readability values and complexities. Samples of reviews presented on the platform are, “*Amazing group of people to work with; they will work with you to solve problems and succeed in making great commission. Regularly show appreciation and reward great performance...*” [1] and “*... Worst Business Model I have ever seen. Abusive environment. Cancer Causing Culture. 70 hours per week...*” [2]. These reviews illustrate the wide range of sentiments and topics discussed on the

platform. First, the results from the pre-processing analyses, the determination of the number of topics appropriate for the dataset, as well as the results from the interrater agreement and topic modelling analyses will be outlined. Research questions 2 and 3, concerning the type of topics extracted and their main effects on helpfulness ratings are then assessed, with their covariates. This is then followed by an assessment of hypothesis 2, which evaluated the effects of the topics' interactions with review sentiment on helpfulness ratings. We followed and, where necessary, adapted guidelines as outlined by Banks et.al. (2018) for best practices when pre-processing text and automating text analysis.

Pre-Processing

This process first involved assessing the reviews' Flesch-Kincaid scores which evaluate the extent to which a chunk of text is readable/interpretable. The mean readability was approximately 60, indicating average reading ease for the reviews. There was an average of 42 words and 224 characters per review ranging from wordcounts of 0 (which were then omitted from the dataset) to 2269. A corpus of 24,654 documents (i.e., actual reviews), 344,604 tokens (i.e., database of words, sub-words, and characters etc.) and 7,701 terms (i.e., words), was created after removing "English" and customized stop words (words that aid with grammatical structure and sentiments), stemming the text, and omitting highly frequent and infrequent terms (7280 of 14981 terms were omitted from analyses – to refine the dataset). While we examined reviews that were on extreme ends of the word count spectrum, we did not delete reviews on the basis of length. The extensive dataset represents a variety of reviews which mitigates the effects of certain extremes. Table 5 shows the top 20 words in the sample, across the entire set of reviews, and two holdout samples, after refining the dataset; and the word cloud shown in Figure 4, illustrates the most relevant words in the dataset, as measured using their term-frequency-

inverse document frequencies (TF-IDF; weighted words to assess relevance alongside prevalence), where those increasing in size indicate greater prevalence. Specifically, weighted words such as management, people, pay, benefit, employee, environment, and team stand out as most relevant and frequent in the dataset.

Employee status was entered as a covariate during this stage to assess how the type of word usage differs between former and current employees (as depicted in Figure 5). It appears that current employees tend to use words concerning management, growth opportunities, cultural values, and benefits. The word, “COVID” is also associated with this group, which means that this is a potential present point of discussion for current employees. Less dominant words are associated with former employees, but these terms seemingly revolve around schedule, and compensation. Further, Figure 6 shows a word cluster plot that highlights which words are often used in conjunction with each other. This is helpful in contextualizing the results from the topic modelling analysis.

Table 5

Comparing the top words that emerged from preliminary analyses across the entire sample of reviews, and holdout samples.

Ranking of words	Entire Sample	Holdout Sample 1	Holdout Sample 2
1	work	work	work
2	manag*	manag*	manag*
3	compani*	compani*	compani*
4	peopl*	peopl*	peopl*
5	pay*	pay*	pay*
6	employe*	employe*	employe*
7	benefit	benefit	benefit
8	hour	hour	hour
9	environ*	environ*	environ*
10	team	team	team
11	opportun*	opportun*	opportun*
12	cultur*	cultur*	cultur*
13	place	place	place
14	flexibl*	flexibl*	flexibl*
15	custom*	custom*	custom*
16	train*	train*	life
17	life	balanc*	chang*
18	chang*	life	train
19	balanc*	chang*	balanc*
20	friend*	friend*	friend*

Note. * Indicates that the word is stemmed. *N* for entire sample: 24,687; *N* for holdout sample 1: 12,240; *N* for holdout sample 2: 12,447.

Search for k (number of topics)

To determine the number of topics appropriate for this dataset, we ran an extensive series of possible topic models, in increments of 5, that ranged from 5 to 65 topics, we also ran topics between those increments (when necessary) in case there was better representation there. Given the scope of the dataset, a general recommendation from the R Core Team (2013) is to examine models that ranged from anywhere between 5 to 60 topics (more than 60 for more extensive

datasets); but we used 65 as the cut-off point. These topic models were measured on a number of metrics such as their overall semantic coherence (i.e., how interpretable the topics were) and held-out likelihood (i.e., the predictive capacity for the topic models), which are the main components to evaluate for topic model quality (Banks et.al., 2018).

Figure 4

Word cloud of most prevalent words across reviews, controlled for frequency.

Note. Increasing size of words directly corresponds to prevalence in dataset.

The general trend observed was that smaller topic models had greater semantic coherence, whereas larger models had higher predictive capacity (albeit a few outliers). Based on these metrics, the 20-topic model demonstrated a balance between semantic coherence and held-out likelihood; however, upon closer observation, the topics within the model were less interpretable than the 25-topic model. In general, we found that the largest set of topic models

were too fine-grained to find coherency, and the smallest topic models lacked certain information shown in those that were more extensive. Therefore, we focused our analyses on the topics that were in the mid-range (i.e., the 15, 20, 25 and 30 topic models). By conducting additional analyses of semantic coherence and held-out likelihood for these four topic models, we found that the 25-topic model displayed the strongest predictive capacity. Though the 25-topic model had a lower semantic coherence than the 15 and 20 topic models, we were able to interpret these topics better than the other models in the analyses. Therefore, we evaluated the 25-topic model as it was more interpretable than other topic models, though it displayed lower statistical semantic coherence than the 20-topic model.

Figure 5

Word cloud of most prevalent words used by current, former employees, and those with non-disclosed employee status.

Note. Increasing size of words directly corresponds to their prevalence for that employee status.

Figure 6

Word Cluster plot of words that co-occur with each other.

Interrater Agreement

Once the number of topics were decided and the topic modelling analyses were conducted, we assessed the content validity ratio for the topics' labels. This was done through a matching exercise with three graduate students, trained in both psychometrics and Industrial/Organizational Psychology. This activity was structured with two columns where column 1 comprised the topic labels and column 2 consisted of the groups of stemmed terms (both high probability and FREX words) that are affiliated with a specific label. We randomized the order of both the labels and grouped terms, then asked the participants to indicate which topic label they believe should be associated with each group of terms. An outline of the exercise used is shown in Appendix A. Fleiss' Kappa was found to be .778 ($p < .001$) where one rater diverged in agreement on the topic labels for Growth Opportunities, Work Shifts/Hours and Vacation and

Overtime. The raters overall proposed that two topics seemed to considerably overlap with each other (Company Culture/Climate and Climate and Pace), and as a result, we combined these topics' proportions to create a single theme. Overall, there is substantial better than chance agreement for the topic labels with their corresponding terms.

Topic Modelling

A 25-structural topic model was computed with employee status entered as a covariate. The terms belonging to the individual topics were categorized based on probability, FREX (frequent and exclusive), lift and score values. The probability and FREX scores informed our labelling of topics and subsequent interpretation, as per the recommendations from Banks et.al. (2018). We also used representative text for each topic to understand the context in which certain terms were used in reviews, thus aiding topic labelling. Because the top ten topics for this model were the most interpretable, and the remaining topics were more difficult to decipher, we restricted evaluation to the more coherent topics. This is important in order to refrain from both misrepresenting the discussions presented on Glassdoor, and erroneously overfitting the dataset.

The most prevalent themes were Company Culture/Climate, Work Life Balance and Compensation and Benefits. As previously mentioned, the raters indicated that Company Culture/Climate was redundant with "Climate and Pace" (also listed in the top ten topics), therefore we combined these topics under the umbrella of "Company Culture/Climate", reducing our evaluation to nine topics. The complete list of the 9 topics is presented in Table 6, with their corresponding topic proportions, labels, and example comments. Examining the text for all nine topics showed that the top three topics seemed to dominate discussion points, as there was clear preoccupation with evaluating the nature of the company climate and benefits provided, and how well the job afforded a clear distinction between personal and professional life. By further

splitting the sample in half in a random fashion, we found that both split samples as well as the original sample, had identical words listed as most prevalent in the dataset (see Table 5). This suggests that the type of discussions (i.e., topics) on these platforms tend to be consistent across several reviews. Therefore, it is likely that there are general concerns or praises with respect to the nine topics outlined across thousands of Glassdoor TPORs.

Table 6

Topics with their labels, terms, example comments and proportions.

Topic Labels	Terms	Example Comments	Topic Proportions
Company Culture/Climate	learn, need, experi*, help, grow, differ, fast, cultur*, career, corpor, competit*, slow, challeng*, structur*	They communicate a lot - but it is all downward - there is little tolerance or appreciation for upward communication. Extremely hierarchical organization. The culture is one of aggressive arrogance by the core management team.	11.3%
Work Life Balance	work, people, life, balanc*, friend, stress, workload, hectic	The company is struggling to strike a balance between the need to deliver on some big goals that we have been promising for a long time, and the desire to continue being the small company that takes "work life balance" and family as a top priority.	9.7%
Compensation and Benefits	pay, benefit, salari*, cowork*, posit, health, rais*, 401k, flight, flex	Great flight benefits, medical and life insurance. The pay needs to be higher.	8.0%
Management and Employee Well-being	manag*, employe*, care, upper, chang*, treat, provid*, listen, overwork	They don't care about the employees just as long as the trains moving. Bad management.	6.2%

Training and Support	team, train*, suppo se*, communic*, o rgan, compens*, i mprov*, program* , organ*, will*	Wonderfully supportive program, develops members, encourages inclusiveness, internally non-competitive environment, with an open-door style of communication.	4.9%
Growth Opportunities	environ*, opportun*, growth, advanc*, easi*, room, home	Plenty of opportunities for movement within organization, both upwards and between teams/departments	4.9%
Work Shifts/Hours	hour, shift, free, sc hedul*, food, brea k, lunch	Get a lot of hours during peak season Don't ge t many hours in off season	4.8%
Leadership Quality	leadership, product, focus, direct, senior, chang*, leader, strategi*, vision, smart	Great mid level managers and directors, great product. Poor executive leadership. New CE O doesn't listen or pay attention to what's bein g said or never understands the topics present ed. Too many VPs.	4.7%
Vacation and Overtime	hour, money, paid, vacat*, overtim*, holiday, call, student, mandatori*, school, teacher, weekend	...paid vacation, health insurance Early hours, start times are different everyday, ending times depend on backhauls. No overtime pay, salary based on hourly rate but not paid hourly.	4.5%

Note. *indicates the term is stemmed. Topic proportions refer to the percentage of the entire database of reviews, that the topic is prevalent in.

Upon examining the reviews that best represented each topic, we noticed several overlaps in the themes discussed within that topic. For instance, though review A may present the highest probability of belonging to topic 1, the issues discussed in review A are diverse enough to apply moderately well to the other topics (in other words, the issues discussed in a specific review were not disproportionate enough to determine a review to be related to one topic instead of another).

The nature of the STM is that all documents are assigned varying probabilities for all topics, however, we found that the reviews were inherently varied with respect to the issues discussed. We therefore used the representative text, alongside the word cluster plot (shown in Figure 6) to decipher the context in which the terms were used, rather than examining the text for the proportion of the topic discussed. Of note however, we found that the representative text loaded better for their respective topics, than the 15, 20 and 30 topic models.

As aforementioned, Table 6 displays all topic names with their descriptions, featured words and example review comments, where the topics are ranked by topic prevalence in descending order. The terms shown are those with the highest probability of belonging to that topic. In answering Research Question 2, the topics extracted from the STM analysis are as follows:

Company Culture/Climate. This topic features words that relate to the pace, structure, and inclusiveness of the work environment as well as the company's cultural values. Therefore, reviewers discussed how they experience the environment of work, which leads to an overall commentary on the company's climate and cultural values.

Work Life Balance. This grouping appears to be associated with concerns about the workload, especially as it relates to maintaining a clear boundary between personal and professional lives. Example comments for this topic tend to be abrupt in the description of this facet of the organization, where reviewers generally limited their descriptions of Work Life Balance to "good" or "bad". This topic also appears to be closely related to Company Culture/Climate as reviewers describe a specific attribute about the work environment.

Compensation and Benefits. Words that are consistent with salary, pay increase and benefits, highly load onto this topic. Reviewers tend to give an overview of the company's benefits and compensation, and their level of satisfaction with these offerings. Like Work Life Balance, they seem to not go in much detail about the implications of these benefit packages/compensation plans and the specific areas where they fall short or are substantial.

Management and Employee Well-being. This illustrates words that depict how management performs in the company, especially as it relates to their treatment of employees (indicated by words such as “treat” and “listen”, once examined in context). The cluster plot and the sentiment analysis (discussed in detail further on) illustrate that the manag* stem was highly associated with employees, showing that management is typically discussed in conjunction with how they relate to employees.

Training and Support. This featured words that related to the quality of and/or the extent to which the company invests in improving the skillset of employees. This topic also addresses the company's extension of support (which could be socioemotional) to their incumbents. Therefore, Training and Support refers to the organization's overall provision of resources to their employees – whether for personal development or to accomplish tasks.

Growth Opportunities. Similar to Training and Support, Growth Opportunities also shows terms that relate to the extent to which the organization facilitates self-development, specifically through opportunities for growth and advancement in the company. Therefore, the words used in the topic mainly concerned how reviewers perceived the longevity of their careers in that specific company.

Work Shifts/Hours. This topic displays words that related to the hours and shifts the organization allotted to employees, which is a critical concern for substantial earning. This topic also addresses employees' experiences working on these shifts, as well as their attitudes towards the allotment of hours. The terms representing these topics also seem to be specifically tied to those working in the service industry, but there is a diversity in the representative texts shown for this topic (not all reviews emerged solely from service industries). This topic is conceptually close to Compensation and Benefits.

Leadership Quality. Like Management and Employee Well-being, Leadership Quality also highlighted words that concerned the management, however these words were specifically related to how well managers, and especially higher-level executives lead the company. Thus, words such as “vision” and “strategy” were cited in relation to this topic.

Vacation and Overtime. At face value, this topic relates closely to Compensation and Benefits as well as Work Life Balance. But this topic specifically features words that are associated with how well the organization facilitates a fluidity in work – whether the opportunity is there to earn more (i.e., via overtime) or cut back (i.e., via vacation).

Topic Clusters and Correlations. Overall, there were three topic clusters identified within the 25-topic model through cluster/correlational analyses, where only correlations of at least .3 were shown between topics. Cluster 1 includes the topics Company Climate, Management and Employee Well-being and Work Life Balance, Cluster 2 comprises Compensation and Benefits, Vacation and Overtime and Leadership Quality while Cluster 3 contains Company Culture, Training and Support and Work shifts/Hours. Cluster 1 had the highest prevalence followed by Clusters 2 and 3. Upon closer observation, it appears that Cluster 1 seems to aggregate around the workplace environment; Cluster 2 mainly identifies the

allowances afforded by the company (i.e., vacation, overtime, pay, benefits); and while the third cluster is more difficult to decipher, it appears to mainly highlight immaterial company support as depicted through training as well as hours and shifts allotted by the organization. Apart from the clusters, the two topics (in the top 9 prevalent topics) that were highest in their association with each other were Growth Opportunities and Work Life Balance – both being employee centred, highlight the professional and personal well-being of the employee. A full correlational matrix of the topics can be found in Table 7.

Table 7

Correlational matrix of helpfulness ratings, review sentiment, and the topics' proportions (N =24654)

Variable	M	SD	1	2	3	4	5	6	7	8	9	10
1. HR	0.81	6.65	-									
2. RS	3.54	1.30	-.11	-								
3. WSH	0.05	0.06	-.05	.00	-							
4. WLB	0.10	0.05	-.08	.42	.00	-						
5. V&O	0.05	0.04	-.05	-.11	.26	-.01	-					
6. CC	0.11	0.06	-.08	.51	-.39	.18	-.43	-				
7. GO	0.05	0.03	-.13	.51	-.02	.66	-.04	.35	-			
8. LQ	0.05	0.04	.05	-.00	-.38	-.23	-.43	.35	-.20	-		
9. T&S	0.05	0.03	-.06	.39	-.34	.03	-.31	.50	.16	.28	-	
10. M&E	0.06	0.03	.00	-.40	-.09	-.22	-.04	-.28	-.23	-.02	-.04	-
11. CB	0.08	0.05	-.09	.02	.09	.21	.35	.19	.34	-.35	-.24	.08

Note. Figures in bold show a significant correlation. HR: Helpfulness Ratings; RS: Review

Sentiment; WSH: Work Shifts/Hours; WLB: Work Life Balance; V&O: Vacation and Overtime;

CC: Company Culture/Climate; GO: Growth Opportunities; LQ: Leadership Quality; T&S: Training and Support; M&E: Management and Employee Wellbeing; CB: Compensation and Benefits.

Though the topic cluster analyses identified three groups among the topics, the themes seem to generally aggregate around two higher order categories – Work Environment (i.e., Company Culture/Climate, Work Life Balance, Management and Employee Well-being, Leadership Quality and Growth Opportunities) and Company Offerings (i.e., Compensation and Benefits, Training and Support, Work Shifts/Hours and Vacation and Overtime). Work Environment concerns descriptions of the nature of the workplace – how employees experience the company’s climate, and the treatment and the quality of management/leadership that contribute to the company’s overall ambiance. Company Offerings address topics that concern the provisions that the organization offers to incumbents, whether through opportunities to earn more (for example, via overtime or hours) and/or enjoy tangible/intangible benefits (for example, via benefits, vacation, or socioemotional support).

Covariate Effects. The STM was then used to derive the nature of the difference between the topics discussed by former and current employees, for exploratory purposes. As shown in Figure 7, employee status did not dictate differences in how much of the discussion was predominated with Vacation and Overtime. However, among the nine topics examined in this study, former employees discussed six of these topics in slightly higher prevalence than current employees. Specifically, current workers discussed Work Shifts/Hours at a higher rate, and former incumbents spoke more about Work Life Balance. Both types of workers discussed the company offerings and the nature of the work environment in an equivalent manner, but it is apparent here that former employees address a wider range of topics.

Figure 7

Effect of Employee Status on Topic Prevalence.

Note. Topics with numbers are less prevalent, and not labelled.

Sentiment Analysis

Before submitting the document-topic proportions for hierarchical regression analyses, we examined if individual ratings could serve as a proxy variable for review sentiment. These regressions would inform which topics predict helpfulness ratings and assess if these effects vary with sentiment (as indexed by individual ratings). However, the underlying assumption with using individual ratings to represent sentiment is that a lower rating is associated with displeasure, or a negative attitude toward the company. Though this may be the case, it is possible that a lower rating is not necessarily associated with a negative sentiment but could indicate that the reviewer wants to see certain improvements in the company, without ill-feelings about those specific flaws. For instance, “organization needs to offer a more diverse benefits package for employees with families” versus, “I am frustrated with the organization’s benefits

plan for employees with families” shows that the same flaw is highlighted about the company, but the latter is associated with an explicit negative sentiment. Therefore, it is not always the case that a lower individual rating would be associated with a negative attitude.

To address this potential confound, we conducted a sentiment analysis of the data using the Afinn (Nielsen, 2011), Bing (Bing, 2010) and Syuzhet (Naldi, 2019; created by Matthew Jockers) dictionaries. These dictionaries include a very large database of emotive words with assigned sentiment scores or codes, where the Syuzhet dictionary provides access to the sentiment tool developed by Natural Language Processing researchers at Stanford university. Once we computed this analysis, we found small to moderate positive correlations between these dictionaries and individual ratings (see correlational matrix in Table 8). This affirms that there is an association between the sentiments of the review and individual ratings, such that as individual ratings increase, so do positive sentiments. However, while we had previously found in part one of this study that individual ratings had a moderate association with helpfulness ratings, we found that the dictionaries had very weak associations with this outcome. This illustrates that while individual ratings may be an appropriate indicator of review sentiment, they could also operate as an index of other cues, external to sentiment.

Predictions

Main Effects. Review sentiments were entered in the first step of the hierarchical regression, the document-topic proportions in the second step, and the interaction terms between review sentiment and the proportions in the third step, with helpfulness ratings as the outcome in this model. Considering that previous analyses showed that review sentiments were the strongest predictors of helpfulness ratings, it was important that we controlled for the effects of this variable first. The correlations between the topics’ proportions, individual and helpfulness

ratings are depicted in Table 7, whereas the correlation matrix of the topics' interaction with review sentiment and helpfulness ratings are illustrated in Table 9. As shown in Tables 10 and 11, the third model with interactions predicts a significant change in the variance of helpfulness ratings, showing that the topics' proportions had direct effects on helpfulness ratings, where the sentiments of certain topics contextualize their influence on review helpfulness.

Table 8

Correlation matrix for the sentiment dictionaries and review sentiment (N = 24,570).

Variable	Mean	SD	1	2	3	4
1. Review Sentiment	3.55	1.29	-			
2. AFINN Vector Dictionary	4.06	5.03	.30*	-	-	
3. Bing Vector Dictionary	1.93	2.26	.25*	.79*	-	-
4. Syuzhen Dictionaries	2.05	2.16	.15	.80	.79	-

Note. * Indicates significant correlations at $p < .001$

These findings favourably answer Research Question 3 where the topics' proportions showed varying effects on review helpfulness. Specifically, Growth Opportunities, and Company Culture/Climate had the highest relation to review sentiment, where an increase in the discussion of these topics was associated with increasingly positive sentiments. These topics also predicted the lowest helpfulness ratings. This pattern was not always consistent across topics, as Work Life Balance and Vacation and Overtime were associated with more negative sentiments but still predicted a decline in helpfulness ratings, albeit very weak.

Similar to Growth Opportunities and Company Culture/Climate, majority of the other topics' proportions predicted a decline in helpfulness ratings, with the exception of Leadership

Quality and Training and Support. Leadership Quality instead showed an increase in helpfulness as persons discussed this topic in greater proportion, and Training and Support did not significantly predict helpfulness ratings. Overall, the consistent pattern across most topics is that as the reviews increase in the discussion of a specific topic (i.e., topic proportion), they receive less traction in helpfulness ratings.

Interactions. The interaction effects between each topic's document-topic proportions and review sentiment on helpfulness ratings were used to assess Hypothesis 2. These interactions were significant, except for those with Training and Support, Management and Employee Wellbeing and Compensation and Benefits. For these topics, it appears that review sentiments do not seem to affect their influence on helpfulness ratings. Significant interactions revealed consistent patterns across the topics. Specifically, negative discussions within a topic received the highest helpfulness ratings, but as topic proportions increased, helpfulness ratings decreased substantially for the topic's reviews that were negatively discussed. For instance, Work Life Balance (as shown in Figure 8) was discussed in varying proportions – some reviewers addressed this organizational attribute in low, moderate, and high prevalence in their reviews (i.e., different proportions). When Work Life Balance was discussed with a negative sentiment it was regarded as more helpful than moderate and positive sentiments. However, as the proportion of conversation regarding Work Life Balance increased along with negative sentiments, helpfulness ratings declined. Specifically, low topic proportions associated with negative sentiments were considered more helpful, than high topic proportions with negative sentiments, which fail to support Hypothesis 2. We had hypothesized that an increase in the negative discussion of a topic would have attracted higher helpfulness ratings, but the reverse was found.

Figure 8

Relationship between Work Life Balance's topic proportions moderated by review sentiment

(N=24,539)

Conversely, positive sentiments within TPORs for the topics' document-topic proportions did not show changes in their trends as topic proportion increased. Another notable trend is that for Growth Opportunities (shown in Figure 9), negative reviews were also most helpful when the topic was discussed in low proportion. However, as topic proportion increases, negative reviews were considered least helpful. This is contrary to the other trends where negative reviews experienced a sharp decline but were still most helpful even with that decrease. Finally, Leadership Quality (Figure 10) was the only topic that depicted a positive trend, where topic proportions were associated with an increase in helpfulness ratings; and this was more prevalent for negative discussions for this topic. The remaining interaction plots for the topics' proportions are shown in Appendix B.

Figure 9

Relationship between Growth Opportunities' topic proportions moderated by review sentiment

($N=24,539$).

Figure 10

Relationship between the topic proportions for Leadership Quality moderated by review sentiment (N=24,539).

Table 9

Correlational matrix of helpfulness ratings, review sentiment, and the topics' proportions' interactions with review sentiment (N =24654)

Variables	M	SD	1	2	3	4	5	6	7	8	9	10
1. HR	0.81	6.65	-									
2. RS	3.54	1.30	-.11	-								
Topics x RS												
3. WSH	0.17	0.23	-.06	.25	-							
4. WLB	0.37	0.24	-.09	.73	.18	-						
5. V&O	0.15	0.15	-.07	.28	.33	.22	-					
6. CC	0.44	0.30	-.09	.75	-.12	.52	-.09	-				
7. GO	0.19	0.14	-.12	.73	.16	.80	.20	.59	-			
8. LQ	0.17	0.17	-.01	.35	-.27	.07	-.29	.51	.09	-		
9. T&S	0.19	0.13	-.08	.70	-.09	.42	-.02	.70	.48	.45	-	

10. M&E	0.21	0.10	-.10	.48	.04	.29	.08	.28	.30	.16	.41	-
11. CB	0.29	0.19	-.11	.51	.20	.48	.44	.21	.58	-.12	.16	.32

Note. Figures in bold show a significant correlation. The topics presented are interacting with Review Sentiment. HR: Helpfulness Ratings; RS: Review Sentiment; WSH: Work Shifts/Hours; WLB: Work Life Balance; V&O: Vacation and Overtime; CC: Company Culture/Climate; GO: Growth Opportunities; LQ: Leadership Quality; T&S: Training and Support; M&E: Management and Employee Wellbeing; CB: Compensation and Benefits.

Table 10

OLS Hierarchical Regression Results for Part Two Study Variables (N=24,654)

Model	Independent Variable	β	t	p
1	Review Sentiment	-0.11	-17.22	<.001
2	Review Sentiment	-0.03	-3.73	<.001
	Work Shifts/Hours	-0.10	-13.09	<.001
	Work Life Balance	0.00	-0.50	.614
	Vacation and Overtime	-0.08	-10.71	<.001
	Company Culture/Climate	-0.13	-13.12	<.001
	Growth Opportunities	-0.06	-6.28	<.001
	Leadership Quality	-0.01	-1.07	.284
	Training and Support	-0.05	-5.82	<.001
	Management and Employee Well-being	-0.08	-10.92	<.001
	Compensation/Benefits	-0.06	-8.12	<.001
3	Review Sentiment	-0.12	-3.38	.001
	Work Shifts/Hours	-0.13	-6.44	<.001
	Work Life Balance	0.14	4.38	<.001
	Vacation and Overtime	-0.12	-5.60	<.001
	Company Culture/Climate	-0.22	-6.24	<.001
	Growth Opportunities	-0.28	-7.02	<.001
	Leadership Quality	0.07	3.09	.002
	Training and Support	-0.02	-0.73	.468
	Management and Employee Well-being	-0.07	-4.24	<.001
	Compensation/Benefits	-0.08	-3.53	<.001
	Work Shifts/Hours x RS	0.05	2.18	.030
	Work Life Balance x RS	-0.18	-4.41	<.001

Vacation and Overtime x RS	0.05	2.09	.037
Company Culture/Climate x RS	0.16	3.37	.001
Growth Opportunities x RS	0.28	5.44	<.001
Leadership Quality x RS	-0.08	-3.46	.001
Training and Support x RS	-0.02	-0.56	.578
Management and Employee Well-being x RS	-0.01	-0.44	.661
Compensation/Benefits x RS	0.04	1.59	.111

Note. RS: Review Sentiment

Table 11

Part two summary of omnibus results for simultaneous regression models (N= 24,654)

Model	<i>R</i>	<i>Adjusted R2</i>	Standard Error	F Change	<i>p</i>
1	.11	.01	6.61	296.50	<.001
2	.20	.04	6.52	78.79	<.001
3	.21	.04	6.50	14.77	<.001

Post-Hoc Analyses. After examining the effects of each topic's proportions on helpfulness ratings, as well as their interaction effects on helpfulness, we found that most topics (five of six) trended negatively with helpfulness ratings, though sentiment was controlled for. This calls into question if readers find these topics unhelpful or if they consider the elaboration of a specific topic (with the neglect of discussing other topics) unhelpful. It is also possible that the length of the review (i.e., word count) potentially influences how these topic proportions affect helpfulness ratings – specifically, are lengthy but narrowed (i.e., predominantly discussing one or a few topics) reviews more helpful than lengthy but varied (i.e., discussing several topics) reviews?

We conducted post-hoc analyses to address these questions. Specifically, we took the highest topic proportion loading for each document to represent the extent to which it captured a variety of topics (i.e., the document's level of focus). It is important to recall that each document

is assigned a proportion value for each topic, which all sum to 1. In other words, a review is assessed on the extent to which it captures each topic. If a review's highest loading was .7, and this was associated with Company Culture/Climate, it indicates that 70% of the document contained information related to this topic. It further highlights that the review was more focused on one or a few topics. Conversely, if the highest loading for the same review was .3, it means that the review was very varied in the type of information it contains. Thus, lower proportions indicate that the review addressed a variety of topics.

The highest proportions for each document (i.e., Max Proportion), as well as their interactions with review length (or word count), were regressed on helpfulness ratings. The findings revealed that reviews that were more narrowed in their analysis of a specific topic, received less helpfulness ratings. And, while lengthier reviews were generally considered most helpful, long but varied documents attracted the highest helpfulness ratings (see Figure 8). The regression weights were strongest and negative for Max Proportion suggesting that it is a potential driver of what influences topic proportions in their effects on helpfulness ratings. Therefore, *how* a topic is discussed (length and variety), is as, or more important than *what* is discussed. The correlation matrix, summary for the regression models and the regression results are shown in Tables 12, 13 and 14, respectively. Figure 8 illustrates the interaction between Max Proportions and Review Length on helpfulness votes.

Table 12

Correlation matrix for variables used in post-hoc analyses (N =24,539)

Variable	Mean	SD	1	2	3	4
1. Helpfulness Ratings	0.65	1.85	-			
2. Review Sentiment	3.55	1.29	-.33*	-		
3. Max Proportion	0.16	.06	-.18*	.27*	-	
4. Review Length	36.51	54.45	.40*	-.20*	-.18*	-

Note. * Indicates significant correlation at $p < .001$.

Table 13

OLS Hierarchical Regression Results for post-hoc analyses (N=24,486)

Model	Independent Variable	Coefficient	<i>t</i>	<i>p</i>
1	Review Length	.014	11.56	<.001
2	Review Length	.013	65.04	<.001
	Max Proportion	-3.32	-18.19	<.001
3	Review Length	.014	32.64	.001
	Max Proportion	-2.93	-13.40	<.001
	Review Length x Max Proportion	-.011	-3.23	.001

Note. Outcome variable: helpfulness ratings.

Table 14

Summary of omnibus results for hierarchical regression models for the post-hoc analyses (N=24,486)

Model	<i>R</i>	<i>Adjusted R2</i>	Standard Error	F Change	<i>p</i>
1	.404	.163	1.69	4769.09	<.001
2	.417	.174	1.68	330.83	<.001
3	.418	.174	1.68	10.42	.001

Note. Outcome variable: helpfulness ratings.

Figure 11

Relationship between maximum proportion and helpfulness ratings at different levels of word count ($N = 24,486$).

CHAPTER 4: Discussion

Despite the potentially widescale impact of online TPORs on recruiting images and applicant attraction, the examination of TPORs has been largely unaddressed in the literature (Pernkopf & Wolfgang 2020). Previous studies have identified the need to address this gap within the recruiting domain (Könsgen et.al., 2018; Pernkopf & Wolfgang, 2020), which further supported the current study's efforts to examine the properties of TPORs. First, we evaluated the sentiment of reviews through their individual ratings and how they vary with helpfulness ratings. We then assessed how this relationship is moderated by overall company ratings, the extent to which reviews' sentiments deviate from others, and the reviewers' employee statuses. We also examined how these predictors fare in their impact on helpfulness ratings. After, we proceeded with the topic modelling analysis to retrieve the most prevalent themes discussed in the TPORs; they were examined for their main and interaction effects' (with review sentiment) on helpfulness ratings.

Overview of Main Findings and Theoretical Implications

Part 1 – Peripheral Cues

We expected that negative reviews would increase the likelihood of review helpfulness and found support for this notion. This finding aligns with previous research on OCRs (Online Customer reviews) in the marketing literature where negative reviews are more influential in decision-making and information adoption when purchasing consumer products and services (Filieri, 2015; Filieri et.al., 2018; Filieri et.al., 2019). More broadly, it is consistent with the negativity bias premise where negative information is more influential in decision making than positive information (Vaish et.al., 2008). The SVM and ELM argue that the increased elaboration of and persuasion by messages can occur as a function of being presented with

personally relevant information (Falk and Scholz, 2018; Howe and Krosnick, 2017; Petty & Cacioppo, 1986). Thus, negative information may be perceived as being relayed in the best interests of the audience, as they are consistently regarded as more credible (Cheung et.al., 2009; Lo & Yao, 2019). Combined with the potentially high-stakes nature of pursuing employment intentions with an organization, it is reasonable to conclude that more skeptical information receives higher instances of helpfulness ratings.

Our second hypothesis focused on whether the relationship between individual and helpfulness ratings would be moderated by overall company ratings. For example, would an overall positive average rating from many reviewers mitigate the negative impact of individual negative reviews? Our data demonstrated that the negative relationship between these two variables is reduced when overall company ratings were high. In other words, the capacity for negative reviews to impact review helpfulness and potentially, information adoption, is lowered when the company communicates higher brand equity or image through their overall ratings. These results are important as studies in the recruiting literature that have assessed overall ratings' impact on application attraction had leveraged experimental designs, where consensus scores buffered the effects of negative reviews on applicant attraction (Carpentier & Van Hove, 2020; Stockman et.al., 2020). Considering that helpfulness ratings are affected in a similar pattern as applicant attraction, it is possible that these two constructs are conceptually linked. This further supports the notion that reviews impact important recruiting outcomes (Carpentier & Van Hove, 2020; Stockman et.al., 2020), consequently affecting the quality and depth of applicant pools (Chapman & Webster, 2006; Van Hove & Lievens, 2009). Therefore, this field study corroborates the utility of overall company ratings, and by extension brand image, by illustrating that they also demonstrate capacity to dampen information adoption, through review

helpfulness, of negative reviews through their effects on helpfulness ratings. Further, this finding is consistent with research on the benefits of corporate image and its capacity to mitigate the effects of scandal and negative WOM (Allen et.al., 2007; Filieri, 2015; Melian-Gonzalez and Bulchand-Gidumal, 2016; Yu, 2019). While this may be the case, corporate image was still not able to completely buffer the impact of negative reviews, which suggests that other factors may be at play that can aid with mitigating these effects.

We also examined the extent to which a review's deviation from consensus scores (indexed by overall ratings) moderates the relationship between individual and helpfulness ratings. We had proposed, based on the implications from previous research (Cheung et.al., 2009, Filieri, 2015; Könsgen et.al., 2018; Pernkopf & Wolfgang, 2020), that a review's individual rating that deviated less from the aggregate company score, would be regarded as more helpful. However, we found the reverse in this study where deviating reviews were associated with higher helpfulness ratings. The moderation analysis revealed that negative reviews that deviated more from overall ratings were regarded as more helpful than highly deviating positive reviews. This suggests that negative reviews that go against the grain, are likely to stand out to reviewers and may increase the likelihood of attracting helpfulness ratings. Overall positive company ratings indicate a resilient consensus among the reviewers about the overall experience with working in the organization. Consensus information has been found to be inherently persuasive as it illustrates both consistent imagery about the company in question, and social agreement in attitudes (Deng et.al., 2020; Falk & Scholz, 2018). We anticipated that deviating reviews would communicate unique and personal biases that cannot be generalized to employee experiences (Filieri, 2015; Könsgen et.al., 2018), which would subsequently decrease their credibility.

We see these consensus effects where an increase in overall company ratings was associated with less helpfulness ratings, which suggests that if readers have noted that an organization portrays a strong image, then the motivation to peruse individual reviews subsequently declines. On the contrary, deviating reviews are not indicative of consensus information, thus we expected that reviews with similar sentiments would be more persuasive as they are likely perceived as more credible. However, because deviating negative reviews specifically attracted higher helpfulness, it may be the case that readers are seeking a comprehensive account of the organization that is not overwhelmingly positive. Therefore, they likely find the negative information helpful in acquainting themselves with the nature of the company and understanding instances where they may potentially have a negative experience. However, they may ultimately be persuaded by the consensus in the reviews' sentiments. Either way, it would be important to determine what exactly these deviating reviews are aiding with processing, and to what audience they are most useful for. The current study's findings indicate the naturally occurring but heightened search for negative and/or standout information when perusing employee review TPPs.

We then inquired whether the perceptions of the review source would be influential in processing TPORs. Specifically, would reviews from former employees be seen as more helpful than those from current employees? Overall, we found that readers generally considered information from former employees as more helpful, and this was more salient than current employees, when their reviews were negative. Though former employees were more influential, the plot shows that increasingly positive sentiments are likely considered most helpful from a current employee. Overall, these findings not only suggest that reviewers are attentive to employee status but that this status projects varying effects on review helpfulness at different

levels of sentiment. Thus, reviewers seem to be more interested in, and persuaded by negative information from former workers, and positive information from current employees.

It is likely that credibility plays an important role in deciphering these results. Specifically, the rationale may be consistent with arguments from Saini and Jawahar (2019) where former employees do not have personal stakes in the success and longevity of the company, therefore increasing their willingness to be more transparent in their reviews. Prospective applicants may therefore consider information from former employees to be more reliable, striking a parallel between the persuasive nature of TPPs and organization-mandate web-based employee testimonials. TPPs feature information from anonymous reviewers with no obligation to improve organizations' images (Van Hove & Lievens, 2007). This is akin to former employees who are likely perceived to be no longer affected by their relationship with a company. Therefore, they could be regarded as being more transparent and credible when sharing their experiences, especially if negative.

This is inconsistent with arguments proposed from focus group participants in Pernkopf and Wolfgang's (2020) study, where they were generally less trusting of negative reviews from former employees, as they are likely to be more biased (given perceptions that former employees could be dissatisfied with their discharge). Therefore, when they highlight positive experiences, it would further substantiate their arguments. While the study's findings illustrate otherwise, we still see an important latitude in these results where the effects of employee status are contingent on the reviews' sentiments. Specifically, though former employees were generally more helpful, this was more salient when their reviews were negative. They may be considered as more willing to discuss negative experiences with the company and may be met with skepticism should they present positive information. Further, current employees' reviews were more positively

positioned which suggests that they are more hesitant to highlight the negative experiences of the company. This may be in fear of harming the corporate image and/or the implications of being recognized on these platforms should they share negative experiences. Thus, readers likely perceive current employees as being more willing to present positive information about the company and may mistrust negative sentiments. Further, these results generally align with arguments proposed from the ELM, that persons are perceptive of message source when processing its respective content (Petty & Cacioppo, 1986; O’Keefe, 1990). While message source is typically regarded as a peripheral cue in information processing, it is not clear whether it is primarily being processed centrally or peripherally on employee review TPPs. Given the elaboration that readers likely invest in determining whether former employees can be a trusted source of information (Pernkopf & Wolfgang, 2020) suggests that prospective applicants deeply consider the implications of employee status when evaluating reviews. Additional research should further probe the rationale behind former employees being perceived as more helpful when they share negative experiences; this may guide organizations’ initiatives in developing effective mitigating strategies for negative reviews.

Our final query in the first part of this study involved assessing how the TPOR components of interest, predicted helpfulness ratings in a multiple regression analysis. We raised the question of which predictors in question would have varying effects, thus did not specify which would be more influential. Review sentiment predicted helpfulness ratings the strongest, followed by deviation scores and employee status. Overall organizational ratings were not found to significantly predict helpfulness ratings when included in this model. Therefore, these results demonstrate that sentiment takes precedence in processing reviews among the peripheral cues examined in the current study. Reviews that deviate from the norm (especially those that are

negative) are more impactful than the perceived consensus of attitudes toward the company. These variables also overlap with some of the OCR components Filieri (2015) examined where, outside of information quality, customer ratings had the highest effect on information diagnosticity (i.e., the extent to which assessing information is aided) followed by rank (i.e., overall product ranking) and source credibility (Wang et.al., 2019). Thus, review sentiments were still notably influential for customer reviews, as it has also been in employee reviews. In general, Filieri (2015) uncovered that information quality was the strongest predictor of informational diagnosticity. Though this component was not examined in our study, it leaves significant implications to be considered.

Part 2 – Topic Modelling

The research questions for part two of this study inquired which topics would be discussed on the employee review TPPs and whether these topics would have differential effects on helpfulness ratings. The findings of the current study have presented favourable responses to these inquiries. We also hypothesized that the effects of topic proportions on helpfulness ratings would be more pronounced for negative discussions in higher topic proportions. However, our findings illustrated otherwise where topic proportions' effects on helpfulness ratings were more effective if discussed in lower prevalence with negative sentiments.

Similar to previous studies that have extracted topics from employee TPPs to assess their overlap with job satisfaction indices and corporate performance (Sainju et.al., 2021), our themes also largely addressed the Work Environment (i.e., topics related to Leadership, Growth Opportunities, Work Life Balance) and Company Offerings (i.e., topics related to corporate benefits and support provided). Given that the overarching goal of our topic modelling analyses was to summarize and contextualize the core discussions on these platforms, we believe that the

themes extracted appear to be consistent with Image (Beach 1990; Lievens & Highhouse, 2004) and Fit theories (Cable & Edwards, 2004) from the recruiting literature.

In the qualitative assessment of the two higher-order clusters identified in the text analyses, we observe that they are reminiscent of the symbolic and instrumental images of organizations. The image theory, conceptualized by Beach (1990) and further expanded by Lievens and Highhouse (2004), proposes that persons tend to leverage a dual processing strategy in determining their level of attraction toward a company. This strategy involves examining how they perceive the company's objective attributes (i.e., instrumental image), such as the benefits packages, compensation plans, and other material/immaterial support offered. Further, prospective applicants also tend to assess the symbolic image of the company. This refers to the organizational values that they believe are fundamentally expressed between and among employees, as well as other stakeholders within the work environment. We propose that the instrumental image of a company can be construed through the topics clustered around the Company Offerings theme identified, and that the Work Environment cluster provides valuable insight into the company's symbolic image.

Considering that prospective applicants also assess their fit with organizations, it is expected that they use their overall impressions of the company's instrumental image to infer the latitude of need-supplies fit with the company. Need-supplies fit refers to the assessment of the alignment between what a company provides and the sufficiency of these provisions for the prospective applicant (Peltokorpi et.al., 2019). Perceptions of the instrumental image of the organization, through Company Offerings, is likely to provide insight into whether the company would be able to fulfil a job seeker's personal needs. Additionally, supplementary fit is concerned with the alignment of values (example: cultural, work and career growth values)

between the person and organization (Cable & Edwards, 2004). Thus, in assessing the symbolic image of the company through the Work Environment cluster, one may be able to determine their level of supplementary fit.

We used topic proportions to examine the influence of these topics on helpfulness ratings. While it was found that the topics' proportions had differential effects on review helpfulness, and majority presented an interaction with review sentiment, it is important to consider the context of these findings. Topic proportions were the most appropriate method in the current study to capture the discussions in a single review. We could have utilized the highest topic loading for each document to indicate which topic the review mostly represents, creating a more refined dataset. However, this would assume that most reviews address a dominant theme, which is not the case in the current study. While most topics' proportions illustrated negative regression trends with helpfulness ratings, these findings may instead be reflecting an interaction between the proportions, and not the topics themselves. Therefore, we caution that a negative trend between a topic and its helpfulness ratings, is not necessarily indicative of that topic being unhelpful. Instead, analyses of the variedness of each document illustrate that the more focused the review is, the less helpful it likely is to the reader. And considering that more focused reviews are associated with positive sentiment (see Table 12), this may be contributing to the declines that we observe with topics' proportions and helpfulness ratings. We therefore propose that future studies effectively account for proportion (preferably in a more controlled environment), in examining the effects of specific topics on review helpfulness and/or information adoption.

Against the backdrop of the nuances that need to be observed when interpreting these findings, we notice that topics concerning the symbolic image of the company were collectively

more prevalent than those for instrumental image. This means that reviewers tend to discuss a wider variety of topics related to the Work Environment, than that of the Company Offerings. And while most topic proportions negatively predicted helpfulness ratings (if significant), those concerning the need-supplies fit were collectively weaker in their influence. Previous studies have illustrated that, signals about the work environment and climate, influence organizational attraction more than factors concerning material benefits (Chapman et.al., 2005). Our findings insinuate that discussing cues about the work environment (e.g., Growth Opportunities, Work-Life Balance, Management and Employee Well-being) were more collectively meaningful to reviewers as these may have graver implications for their level of satisfaction with their job/company and overall organizational attraction. Considering that these factors tend to be central in organizational attraction (Chapman et.al., 2005), it is likely that readers may be more attentive to topics concerning supplementary fit, than they are for need-supplies fit.

Another potential rationale for why readers seem to be less attentive to assessing their need-supplies fit is that objective signals about the workplace, such as the compensation and benefits, are generally readily available on several TPPs through an aggregation of reviewers' wages. Thus, prospective applicants may have already processed the company's instrumental image before they read the reviews, or they may have decided to read the reviews because they derived a satisfactory need-supplies fit. This likely facilitates information processing by mitigating the information overload that tends to occur on these platforms (Cao et.al., 2011; Zhu et.al., 2014). Therefore, readers may peruse qualitative accounts to retrieve information about the company that is not readily available through Glassdoor's company's overview. For instance, while Glassdoor may provide certain "tags" (i.e., an overview of the top two pros and cons outlined in reviews) about the company, potentially regarding the work environment, readers

would only be able to unpack the nuances of the work environment through reading the reviews in more detail.

Additionally, the expectations for satisfactory compensation differ across prospective applicants, so it may be more helpful for readers to assess objective data on that facet. For instance, a reviewer and a prospective applicant may have differing criteria of what constitutes satisfactory compensation, which renders discussions about benefits/compensation packages as unhelpful. To that point, this could be one of the factors that prompts reviewers to focus on more environmental attributes in their accounts, and for readers to be less attentive to cues related to the company's instrumental image.

While it is crucial to consider the differential effects of topics on review helpfulness, it is also important to assess *how* these themes are discussed in reviews. The general trend that was noticed across most of the topics that significantly interacted with review sentiment, was that there was a general decline in helpfulness ratings as the topic proportion increases. This may be indicative of oversharing on a specific topic, which is likely not helpful to readers who want to know about other company attributes, or a variety of job-organizational characteristics. Aligning with ELM, the number of arguments presented in a given message can also be processed peripherally, as they further substantiate the credibility of the overarching information shared (Petty & Cacioppo, 1986). A reader may specifically want to learn about the global leadership of the company; however, a review may instead largely focus on lower-level management which would not be considered helpful to that prospective applicant. Discussing a specific facet of an organization is likely helpful if done in a moderate amount, as reviewing a larger variety of topics could increase the likelihood that readers may find a discussion point that is helpful to them or increase perceived credibility (Petty & Cacioppo, 1986; Petty & Wegener, 1998). This is

also consistent with SVM where the audience is more perceptive to and persuaded by information that is consequential to them (Falk & Scholz, 2018). On the other hand, while readers may want to learn more about a specific topic, an extended dialogue about this issue may eventually surpass the point of providing needed information. In other words, readers could have already been adequately informed about a discussion point that is personally relevant to them, even before completely reading the review about that specific topic.

Within each topic, negative sentiments were associated with higher helpfulness ratings and moderate to positive sentiments showed little to no change in their trends as the topic's proportion increased. Though negative sentiments have been previously shown to gain more traction, it appears that when they interact with topic proportions, there is a limit in how influential this sentiment is, as the trend eventually declines. Readers may be more convinced by, and/or seek out negative information to understand worst case scenarios of being employed in the organization, however an excessively negative account about a specific job-organizational characteristic may be regarded as unbalanced. Further Filieri (2015) showed that an inordinate amount of negative sentiment associated with a review was considered less helpful, and that there are important factors that moderate their effects. In this case, it appears that if a topic is discussed in a low proportion, with negative sentiments then it is considered most helpful.

Negative reviews are generally described as being less biased, through being perceived as more credible accounts (Cheung et.al., 2009; Lo & Yao, 2019) of company experiences. However, it is possible that a more negative account of a specific topic reaches a point of mistrust about the credibility of the TPOR shared. Readers may potentially call into question the credibility of these highly emotional, and excessively negative accounts (Pernkopf & Wolfgang 2020). Previous research has shown that extremely negative reviews are not always helpful

(Filieri, 2015; Pernkopf & Wolfgang, 2020). And as mentioned, more focused reviews tend to be considered less helpful than a varied account. It is likely that while an overall negative review may be regarded influential, a specific topic addressed within the review is effective if it is discussed in low proportions. And if it is discussed in lower proportions, it is most helpful if accommodated with negative sentiments, likely facilitating information diagnosticity and information adoption, as smaller chunks of information with an unambiguous negative sentiment (Filieri, 2015), is potentially easier to process. However, if the topic is discussed in a larger proportion and is also associated with negative sentiments, then this likely challenges the focus of the review, and its persuasive quality. This is due to the possible perception that negative elaborations may be regarded as overly focused rants that are less balanced in their accounts. Considering the speculative nature of these findings, future research could explore the limited effects of sentiments on persuasion, and if elaboration on a specific attitude could operate as a moderator.

Growth Opportunities, Company Culture/Climate and Work Life Balance's interaction with review sentiment were noticeably strong predictors of declines in helpfulness ratings. These topics were also discussed with an overall positive sentiment, that was found to be more salient than other topics. This potentially explains the rationale behind their strong, but negative association with helpfulness ratings, especially as the topic proportion increases. It may be reflective of the seriousness that readers treat environmental variables, especially those concerning opportunities for growth, and the organization's cultural values. Therefore, readers may find it more pertinent to receive a nuanced and balanced (not only positive) account of these topics, that when they do not, it more severely penalizes the reviews' helpfulness.

These findings collectively present implications for what issues are most salient for current/former employees and what factors prospective applicants find helpful when perusing online reviews. Considering that helpfulness ratings are affected by TPOR components in a similar fashion as applicant attraction, suggests that helpfulness ratings could be used to discern how attraction is ultimately affected on TPPs. Therefore, they further highlight the need for organizations to be more attentive to their online portrayal, where it would be in their best interests to address the various factors that likely affect their recruiting images.

Practical Implications

We were cautious with not conflating helpfulness ratings with persuasion but acknowledge that they could be indicative of information adoption (Filieri et.al., 2015; Zhu et.al., 2014). This study further substantiates the need for companies to heed how they are portrayed on these online TPPs, especially when this area of employer review research is largely unexplored. Considering that negative reviews, especially those deviating from the norm, are associated with higher helpfulness ratings, implies the kind of information that is likely to be adopted on these platforms. If companies neglect to address how their image is being portrayed through their overall ratings, this can facilitate the ease at which negative information is adopted and used for decision-making. Thus, it is important that organizations (irrespective of size) develop an aptitude for considering how they are depicted online and develop strategies to address/mitigate this impact.

Further, organizations with a low number of reviews have been shown to be less attractive to consumers, and perceptions of company popularity increases with number of reviews (Filieri & McLeay, 2013). If companies create an organizational culture that encourages current and former employees to share their experiences on TPPs, then this may boost their

online profiles, and applicant attractiveness. We expect that this implication would be more useful for smaller companies that are confident in their material and immaterial provisions to employees.

These findings could be leveraged by Human Resource (HR) Management, and by extension, the organization, to understand the main issues that employees face (current and former alike). Specifically, the organization may find it beneficial to participate in routine check-ins with staff about the environment of the workplace. Considering that employees have a voice through TPPs, that likely impacts the organization's corporate image and potential to attract competitive applicants, it is important that companies consider how their climates are experienced by employees. To make these conversations more effective, companies could focus these discussions on the specific issues that employees have raised on Glassdoor, or other TPPs. Companies could therefore benefit from the LDA methods employed in the current study to help them focus on the specific concerns identified, and/or leverage additional resources to uncover issues unique to their organizations. To that point, an important takeaway from these findings is that former employees' reviews attracted higher helpfulness ratings than current employees, and this was even more salient when former employee reviews were moderately to extremely negative. This suggests that readers perceive negative information from former employees as more credible or compelling than reviews from current employees. Further, the LDA analysis inadvertently show that former employees addressed more topics in their reviews, which implies that they outline more arguments in the information shared; thus, increasing the message's persuasive quality (Petty & Cacioppo, 1986). Considering that research from both the marketing and recruiting literatures have outlined potential organizational responses that can mitigate the effects of negative reviews (Carpentier & Van Hove, 2020; Lee & Cranage, 2014), organizations

that actively respond to these reviews on Glassdoor may find it more effective to prioritize those of former employees. Furthermore, HR should consider improving the mechanisms underlying employees' departure from the organization. Given that their negative opinions are potentially more detrimental toward online recruiting images, it is important to examine and address their levels of concerns as they exit the company.

These findings provide insight into what issues are potentially top-of-mind for prospective applicants. Specifically, prospective incumbents may have received negative signals about the company's work environment through these employee TPPs, which may influence their willingness to accept an offer from another competitor. Recruiters and/or interviewers need to be cognizant of how to best address these potential concerns during the recruiting/selection stages which might be critical for increasing the chances of hiring the best candidates for positions. We previously discussed that, applicants who are more competitive, are indeed more sensitive to negative information about the organization (Chapman & Webster, 2006). These applicants are potentially perceptive to cues about the work environment, thus recruiters/interviewers may need to consider ways in which they can initiate discussions about these specific issues, to alleviate or clarify expectations about potential concerns.

Further, the topics highlighted with the highest prevalence in TPORs, as well as their interaction with review sentiment, show varying effects on helpfulness ratings. These illustrate the specific issues that tend to be highlighted by reviewers and assessed by readers, which can then be used to improve the interface of employer TPPs. Specifically, one of the filters available for sifting through reviews is "helpfulness ratings" which organizes the reviews from most to least helpful. Other potential ideas for filters include "Narrative Helpfulness" which would potentially sort reviews with topics that discuss environmental factors as more easily accessible,

than topics that are considered less qualitatively helpful. An extension of this proposition could include an option where users are provided a list of the topics uncovered in this analysis and are asked to check off the topic that they would like to learn more about in reviews. This filter would likely direct readers to reviews that used those topics as key words. These suggestions may potentially enhance users' experiences with TPPs, but also could inadvertently encourage readers to be more scrutinous of an organization (as they would have easier access to information that matters most for their decision-making).

Limitations and Future Research

This study examined TPOR components and helpfulness ratings for thousands of reviews, which presented unique limitations. For instance, depending on the size of the company, helpfulness ratings for reviews could range from 0 to 1000, where helpfulness ratings of 0 were most dominant in this dataset, therefore creating instances of skewed data, extreme outliers, and a potentially restricted range. Thus, examining data in this domain presents challenges with achieving a normal distribution which admittedly, tends to be the case when examining organic field data, where certain parameters are out of researchers' control.

Moreover, the Glassdoor platform presents prompts for discussing both the pros and the cons of the organization, and we combined both into one response for the reviewer. We speculate that these prompts encourage reviewers to consider and share a broader range of experiences with the company, rather than those that were top-of-mind. This structure potentially benefits prospective applicants who are likely to see a more balanced account of different facets of the company, however it complicates the automated text analyses. Specifically, it appears that reviews tend to discuss a wide range of topics, instead of discussing topics disproportionately. Therefore, when examining the most representative texts for each topic, they seemingly appear

to apply to other topics as well. While the LDA was still able to retrieve the consistent topics discussed across reviews, even though the reviews inherently tend to be less reflective of a dominant topic/theme, the process of predicting outcomes is constrained. This is contrary to TPPs such as Indeed.ca, which allows reviewers to share their unprompted top-of-mind experiences, thus increasing the chance that reviews will reflect topics that are more meaningful for reviewers. Future research may benefit from examining if the current study's findings would be consistent across other TPPs. Moreover, further studies could evaluate if the platform's traffic may underestimate the effects shown in this study. Considering that Indeed.ca attracts 250 million unique monthly visits on average, compared to Glassdoor's 60-70 million unique visits, Indeed's reviews may gain higher helpfulness ratings which could alleviate range restriction issues. Indeed also has a "not helpful" vote for reviews which likely facilitates improvements in interpretation, that would be worth examining in future research.

Future research could also examine if these results would still hold in variations of experimental designs – especially the findings concerning sentiment deviation, which were contrary to our expectations. Additionally, these experimental studies could examine the influence of the more prevalent topics on helpfulness ratings, as well as their interaction with review sentiment, to further assess the reliability and validity of these findings. It would also be beneficial for additional research to address the need to determine the relationship between review helpfulness and other important outcomes such as information adoption, persuasiveness, attraction, employment intentions, job pursuit, and applicant attraction. This would serve to further conceptualize and contextualize how we discuss helpfulness ratings in the organizational review context.

Moreover, we also anticipate that these results may vary if certain nested groups are accounted for. For example, these findings may change depending on employee status, companies and TPPs (if examining data across different platforms). Because this study assessed thousands of data points, we expect that individual company variance may be dampened and multilevel modelling (MLM) would therefore be a redundant process. However, we still believe that future research should consider validating the findings of this study using an MLM approach.

Likewise, another limitation is the potential selection bias associated with the companies chosen to extract reviews from. Specifically, less established companies may have less representation in the dataset, considering that Wikipedia likely features organizations with a digital paper trail. The scope of this study's dataset, however, includes reviews from small to mid-size/large companies (as potentially indexed by the number of reviews each company has) and is theoretically extensive enough to drown out unique company variance. We also randomly sampled 450 companies from diverse industries to mitigate these biases. Regardless of these strategies, future research that attempts to employ a similar methodology, may benefit from collecting a larger sample of reviews, and account for industry and company size factors in validating the findings of this study.

Finally, we also anticipate that future research may deem it useful to examine how the quality of reviews (which could be inferred through readability, grammatical structure, review depth, breadth etc.) influence helpfulness ratings, when compared to review sentiment. Considering that examining job reviews carries more weight than those of examining products' (Pernkopf & Wolfgang, 2020), it may be the case that peripheral cues such as sentiment, overpower the influence of information quality. Additionally, it would also be useful to compare

product and employer review platforms and their differential effects on helpfulness ratings and applicant attraction. This may provide a more comprehensive understanding of the underlying motives when perusing these different channels, and how they differentially impact organizations' recruiting images.

Conclusion

Considering the potential repercussions of TPORs on corporate image and organizational/applicant attraction, it is concerning that their impact has been largely overlooked in the recruiting literature. This study aimed to further the discussion of how recruiting continues to evolve and form its own entity outside of organizational control. We examined review helpfulness as an outcome variable that can be understood as an indicator of information adoption. It was found that negatively valenced reviews that were provided by former employees and deviated from the norm, were the most influential in predicting whether information would be considered helpful. We also observed that topics discussed generally concerned either the organization's instrumental or symbolic image, where negative reviews about these topics received higher helpfulness ratings (though only to a certain degree). However, corporate image depicted the ability to partially buffer against negative reviews. Therefore, this research implies that if organizations attend to how they are portrayed on these TPPs, they may be able to counter some of these effects. Given the novelty of this research in the recruiting arena, it would be fruitful to examine what exactly are the implications of receiving higher helpfulness ratings, and how these might affect how companies manage their recruiting practices. This would be crucial in determining what the far-reaching consequences are of TPORs and employee voice, and how companies can adapt to this lack of agency in their recruiting practices.

References

- Allen, D., Otondo, R. & Mahto, R. (2007). Web-based recruitment: Effects of information, organizational brand, and attitudes toward a web site on applicant attraction. *Journal of Applied Psychology*, 92(6), 1696-1708.
- Banks, G., Woznyj, H., Wesslen, R. & Ross, R. (2018). A review of best practice recommendations for text analysis in R (and a User-Friendly App). *Journal of Business and Psychology*, 33, 445-459.
- Bassig (2018). Yelp factsheet: Stats your business needs to know. *Review Trackers*. Retrieved from: <https://www.reviewtrackers.com/blog/yelp-factsheet/>
- Baum, M., Schafer & Kabst, R. (2016). Modeling the impact of advertisement-image congruity on applicant attraction. *Human Resource Management*, 55(1), 7-24.
- Baumer, E., Mimno, D., Guha, S., Quan, E. & Gay, G. (2017). Comparing grounded theory and topic modeling: Extreme divergence or unlikely convergence? *Journal of the Association for Information Science and Technology*, 68(6), 1397-1410.
- Bing, L. "[Sentiment Analysis and Subjectivity.](#)" Invited Chapter for the *Handbook of Natural Language Processing*, Second Edition. March, 2010.
- Buenano-Fernandez, Gonzalez, M., Gil, D. & Lujan-Mora, S. (2020). Text mining of open-ended

- questions in self-assessment of university teachers: An LDA topic modeling approach. IEEE Access. PP. 1-1. 10.1109/ACCESS.2020.2974983.
- Cable, D. M., & Judge, T. A. (2003). Managers' upward influence tactic strategies: The role of manager personality and supervisor leadership style. *Journal of Organizational Behavior*, 24(2), 197–214. <https://doi.org/10.1002/job.183>
- Cable, D. M., & Turban, D. B. (2003). The value of organizational reputation in the recruitment context: A brand-equity perspective. *Journal of Applied Social Psychology*, 33(11), 2244–2266. <https://doi.org/10.1111/j.1559-1816.2003.tb01883.x>
- Cao, Q., Duan, W., & Gan, Q. (2011). Exploring determinants of voting for the “helpfulness” of online user reviews: A text mining approach. *Decision Support Systems*, 50(2), 511–521. <https://doi.org/10.1016/j.dss.2010.11.009>
- Carpentier & Van Hove (2020). Managing organization attractiveness after a negative employer review: Company response strategies and review consensus. *European Journal of Work and Organizational Psychology*
- Chapman, D. & Gödöllei, A. (2017). E-Recruiting: Using technology to attract job applicants. In G.Hertel, D. Stone, R. Johnson & J. Passmore (Eds.), *The Psychology of the Internet at Work Ed*: 211-230. New Jersey: Wiley-Blackwell.
- Chapman, D., Uggerslev, K., Carroll, S., Piasentin, K. & Jones, D. (2005). Applicant attraction

- to organizations and job choice: A meta-analytic review of the correlates of recruiting outcomes. *The Journal of Applied Psychology*. 90. 928-44. 10.1037/0021-9010.90.5.928.
- Chapman, D. & Webster, J. (2006). Toward an integrated model of applicant reactions and job choice. *The International Journal of Human Resource Management*, 17(6), 1032-1057.
doi:10.1080/09585190600696572
- Cheung, M. Y., Luo, C., Sia, C. L., & Chen, H. P. (2009). Credibility of electronic word-of-mouth: Informational and normative determinants of online consumer recommendations. *International Journal of Electronic Commerce*, 13(4), 9-38.
- Cian, L., & Cervai, S. (2014). Under the reputation umbrella. *Corporate Communications*, 19(2), 182–199. <https://doi.org/10.1108/CCIJ-10-2011-0055>
- Deng, W., Yi, M. and Lu, Y. (2020), “Vote or not? How various information cues affect helpfulness voting of online reviews”, *Online Information Review*, 44(4), 787-803.
- Denny, M. & Spirling, A. (2018). Text preprocessing for unsupervised learning: why it matters, when it misleads, and what to do about it. *Political Analysis*. 26. 1-22. 10.1017/pan.2017.44.
- Doldor, E., Wyatt, M. & Silvester, J. (2019). Statesmen or cheerleaders? Using a topic modeling to examine gendered messages in narrative development feedback for leaders. *The Leadership Quarterly*, 1-21.

- Evertz, L., Koliitz, R. & Sub, S. (2019). Electronic word-of-mouth via employer review sites – the effects on organizational attraction. *International Journal of Human Resource Management*, 1-30, DOI: <https://doi.org/10.1080/09585192.2019.1640268>
- Falk, E. & Scholz, C. (2018). Persuasion, influence, and value: Perspectives from communication and social neuroscience. *Annual Review of Psychology*, 69, 329-356.
- Filieri, R. (2015). What makes online reviews helpful? A diagnosticity-adoption framework to explain informational and normative influences in e-WOM. *Journal of Business Research*, 68, 1261-1270.
- Filieri, R., & McLeay, F. (2014). E-WOM and Accommodation: An analysis of the factors that influence travelers' adoption of information from online reviews. *Journal of Travel Research*, 53(1), 44–57. <https://doi.org/10.1177/0047287513481274>
- Filieri, R., Raguseo, E. & Vitari, C. (2018). When are extreme ratings more helpful? Empirical evidence on the moderating effects of review characteristics and product type. *Computers in Human Behavior*, 88, 134-142.
- Filieri, R., Raguseo, E. & Vitari, C. (2019). What moderates the influence of extremely negative ratings? The role of review and reviewer characteristics. *International Journal of Hospitality Management*, 77, 333-341.
- Ghasemaghaei, M., Eslami, P., & Deal, K. & Hassanein, K. (2018). Reviews'

length and sentiment as correlates of online reviews' ratings. *Internet Research*. 28. 00-00. 10.1108/IntR-12-2016-0394.

Glassdoor (2020). 50+ HR and recruiting stats for 2020. *Statistical Reference Guide for Recruiters*. Retrieved from:

<https://www.glassdoor.com/employers/resources/hr-and-recruiting-stats/>

Glassdoor (2021). 40+ stats for companies to keep in mind for 2021. *Statistical Reference Guide for Recruiters*. Retrieved from: <https://www.glassdoor.com/employers/resources/hr-and-recruiting-stats/>

Gregory, C., Meade, A., & Thompson, L. (2013). Understanding internet recruitment via signaling theory and the elaboration likelihood model. *Computers in Human Behavior*, 29(5), 1949-1959.

Hayes, A. F. (2012). PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling [White paper]. Retrieved from <http://www.afhayes.com/>

Highhouse, S., Zickar, M.J., Thorsteinson, T.J., Stierwalt, S.L., & Slaughter, J.E. (1999). Assessing company employment image: An example in the fast food industry. *Personnel Psychology*, 52(1), 151-172, <https://doi.org/10.1111/j.1744-6570.1999.tb01819.x>

Howe, L. & Krosnick, J. (2017). Attitude strength. *Annual Review of Psychology*, 68, 327-351.

Indeed (2020). About Indeed. Retrieved from: <https://ca.indeed.com/about>

Jones, D. A., Shultz, J. W., & Chapman, D. S. (2006). Recruiting through job advertisements:

The effects of cognitive elaboration on decision making. *International Journal of Selection and Assessment*, 14, 167–179. [http:// dx.doi.org/10.1111/j.1468-2389.2006.00342.x](http://dx.doi.org/10.1111/j.1468-2389.2006.00342.x).

Jung, Y., & Suh, Y. (2019). Mining the voice of employees: A text mining approach to identifying and analyzing job satisfaction factors from online employee reviews.

Decision Support Systems, 123(113074), 2-12. <https://doi.org/10.1016/j.dss.2019.113074>

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk.

Econometrica, 47, 263-291.

Könsen, K., Schaarschmidt, M, Ivens, S & Munzel, Andreas (2018). Finding meaning in

contradiction on employee review sites – effects of discrepant online reviews on job applicant intentions. *Journal of Interactive Marketing*, 43, 165-177.

Kruglanski, A. W. (1990). Motivations for judging and knowing: Implications for causal

attribution. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Foundations of social behavior*, Vol. 2, pp. 333–368). The Guilford Press.

Kusumasondjaja, S., Shanka, T. & Marchegiani, C. (2012). Credibility of

online reviews and initial trust: The roles of reviewer’s identity and review valence. *Journal of Vacation Marketing*. 18. 185-195. 10.1177/1356766712449365.

- Lee, C. & Cranage, D. (2014). Toward understanding consumer processing of negative online word-of-mouth communication: The roles of opinion consensus and organizational response strategies. *Journal of Hospitality and Tourism Research*, 38(3), 330-360.
- Lee, J. & Kang, J. (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. *International Conference on Information Systems*.
- Li, S. T., Pham, T. T., & Chuang, H. C. (2019). Do reviewers' words affect predicting their helpfulness ratings? Locating helpful reviewers by linguistics styles. *Information and Management*, 56(1), 28-38. <https://doi.org/10.1016/j.im.2018.06.002>
- Li, Y., Rapkin, B., Atkinson, T., Schofield, E. & Bochner, B. (2019). Leveraging latent Dirichlet allocation in processing free-text personal goals among patients undergoing bladder surgery, *Quality of Life Research*, 28, 1441-1445.
- Liang, F. (2019). The new silk road on Facebook: How China's official media cover and frame a national initiative for global audiences. *Communication and the Public*, 4(4), 261-275.
- Lievens, F., & Highhouse, S. (2003). The relations of instrumental and symbolic attributes to a company's attractiveness as an employer. *Personnel Psychology*, 56(1), 75–102.
<https://doi.org/10.1111/j.1744-6570.2003.tb00144.x>
- Lo, A.S. and Yao, S.S. (2019), "What makes hotel online reviews credible? An investigation of the roles of reviewer expertise, review rating consistency and review

- valence", *International Journal of Contemporary Hospitality Management*, 31(1), 41-60. <https://doi.org/10.1108/IJCHM-10-2017-0671>
- Lopez Hernandez, F., Verdú, E. & Granados, J. & Gonzalez C. (2019). A nondisturbing service to automatically customize notification sending using implicit-feedback. *Scientific Programming. IP*. 1-33. 10.1155/2019/1293194.
- Ludwig, S., Ruyter, K., Friedman, M., Brügger, E., Wetzels, M. & Pfann, G. (2013). More than words: The influence of affective content and linguistic style matches in online reviews on conversion rates. *Journal of Marketing*, 77(1), 87-103.
- Luo, N., Zhuo, Y. & Shon, J. (2016). Employee satisfaction and corporate performance: Mining employee views on Glassdoor.com. In *Proceedings of the 37th International Conference on Information Systems*, (pp. 1-16).
- McFarland, L. A., & Ployhart, R. E. (2015). Social media: A contextual framework to guide research and practice. *Journal of Applied Psychology*, 100(6), 1653-1677.
doi:10.1037/a0039244
- McKenny, A. F., Aguinis, H., Short, J. C., & Anglin, A. H. (2018). What doesn't get measured does exist: Improving the accuracy of computer-aided text analysis. *Journal of Management*, 44(7), 2909–2933. <https://doi.org/10.1177/0149206316657594>
- Melian-Gonzalez, S. & Bulchand-Gidumal, J. (2016). Worker word of mouth on the internet: Influence on human resource image, job seekers and employees. *International Journal*

of Manpower, 47(4), 709-723. DOI: 10.1108/IJM-09-2014-0188

Morrison, Elizabeth W. (2014). Employee voice and silence. *Annual Review of Organizational*

Psychology and Organizational Behavior, 1(1), 173–197.

<https://doi.org/10.1146/annurev-orgpsych-031413-091328>

Naldi, M. (2019). A review of sentiment computation methods with R packages.

Nielsen, F. (2011) A new ANEW: Evaluation of a word list for sentiment analysis in

microblogs. *Proceedings of the ESWC2011 Workshop on 'Making Sense of Microposts':*

Big things come in small packages 718 in CEUR Workshop Proceedings 93-98.

<http://arxiv.org/abs/1103.2903>.

O’Keefe, D. J. (2002). ***Persuasion: Theory and research*** (2nd ed.). Thousand Oaks, CA: Sage

Peltokorpi, V., Bader, A. K., & Froese, F. J. (2019). Foreign and domestic company

attractiveness to host national employees in Japan: A person–organization fit and image

theory perspective. *International Journal of Selection and Assessment*, 27(4), 392–405.

<https://doi.org/10.1111/ijsa.12259>

Pernkopf, M. & Wolfgang (2020). Effects of mixed signals on employer attractiveness: A

mixed-method study based on signalling and convention theory. *Human Resource*

Management Journal. 10.1111/1748-8583.12313.

Petty, R. & Cacioppo (1986). ***Communication and persuasion: Central and peripheral routes***

to attitude change. New York: Springer-Verlag.

Petty, R. E., & Wegener, D. T. (1998). Attitude change: Multiple roles for persuasion variables.

In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social*

psychology (pp. 323–390). McGraw-Hill.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for

Statistical Computing, Vienna, Austria. URL: <http://R-project.org>

Risselada, Hans, de Vries, Lisette, & Verstappen, Mariska. (2018). The impact of social

influence on the perceived helpfulness of online consumer reviews. *European Journal of*

Marketing, 52(3/4), 619–636. <https://doi.org/10.1108/EJM-09-2016-0522>

Rynes, S. L., & Barber, A. E. (1990). Applicant attraction strategies: An organizational

perspective. *The Academy of Management Review*, 15(2), 286–

310. <https://doi.org/10.2307/258158>

Saini, Gordhan K, & Jawahar, I.M. (2019). The influence of employer rankings, employment

experience, and employee characteristics on employer branding as an employer of choice.

Career Development International, 24(7), 636–657. <https://doi.org/10.1108/CDI->

[1120180290](https://doi.org/10.1108/CDI-1120180290)

Sainju, B., Hartwell, C. & Edwards, J. (2021). Job satisfaction and employee turnover

determinants in Fortune 50 companies: Insights from employee reviews from

- Indeed.com, *Decision Support Systems*, <https://doi.org/10.1016/j.dss.2021.113582>.
- Stamolampros, P., Korfiatis, N., Chalvatzis, K., & Buhalis, D. (2019). Job satisfaction and employee turnover determinants in high contact services: Insights from employees' online reviews. *Tourism Management*, 75, 130–147.
<https://doi.org/10.1016/j.tourman.2019.04.030>.
- Stockman, S., Van Hoya, G. & da Motta Veiga, S. (2020). Negative word-of-mouth and applicant attraction: The role of employer brand equity. *Journal of Vocational Behavior*. 118. 103368. 10.1016/j.jvb.2019.103368.
- Sung, K.H. & Lee, M.J. (2015). Do online comments influence the public's attitudes toward an organization? Effects of online comments based on individuals' prior attitudes. *The Journal of Psychology*, 149(4), 325-338.
- Tsai, W. & Yang, W. (2010). Does image matter to different job applicants? The influences of corporate image and applicant individual differences on organizational attractiveness. *International Journal of Selection and Assessment*, 18(1), 48-63.
- Turban, D. B. (2001). Organizational attractiveness as an employer on college campuses: An examination of the applicant population. *Journal of Vocational Behavior*, 58, 293–312.
- Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of Business*, 59, S251-S278.

- Vaish, A., Grossman, T. & Woodward, A. (2008). Not all emotions are created equal: The negativity bias in social-emotional development. *Psychological Bulletin*, 134(4), 383-403.
- Van Hove, G. & Lievens, F. (2007). Investigating web-based recruitment sources: Employee testimonials vs word-of-mouth. *International Journal of Selection and Assessment*, 15, 372 - 382. 10.1111/j.1468-2389.2007.00396.x.
- Van Hove, G. & Lievens, P. (2009). Tapping the grapevine: A closer look at word-of-mouth as a recruitment source. *Journal of Applied Psychology*, 94, 341-352.
- Wang, Y., Wang, J. & Yao, Tang. (2019). What makes a helpful online review? A meta-analysis of review characteristics. *Journal of Electronic Commerce Research*, 19, 10.1007/s10660-018-9310-2.
- Wu, P.F. (2013). In search of negativity bias: An empirical study of perceived helpfulness of online reviews. *Psychology & Marketing*, 30(11), 971–984.
<https://doi.org/10.1002/mar.20660>
- Yu, K. Y. T. (2019). Influencing how one is seen by potential talent: Organizational impression management among recruiting firms. *Journal of Applied Psychology*, 104(7), 888–906. <https://doi.org/10.1037/apl0000386>
- Zhu, L., Yin, G., & He, W. (2014). Is this opinion leader's review useful? Peripheral cues for online review helpfulness. *Journal of Electronic Commerce Research*, 15(4), 267-280.

Appendix A: Matching Labels to Terms Exercise

Matching Topic Labels to corresponding Group of Words (i.e., topic models)

Write the topic model's ***number*** in the box that corresponds to the topic label that you believe best describes it.

Topic Labels		Topic Models
Work Life Balance	<input type="text"/>	1 Key Words: cultur, career, corpor, competit, slow, challeng, structur, mobil, slow, inclus, stabl, upward
Compensation and Benefits	<input type="text"/>	2 Key Words: work, people, life, balance, stress, friend, workload, fo, hectic
Company Culture/Climate	<input type="text"/>	3 Key Words: pay, benefit, salary, cowork, posit, health, rais, 401k, flight, flex
Management and Employee Well-being	<input type="text"/>	4 Key Words: hour, shift, free, schedule, food, break, lunch, pizza, tip, food, free, meal, rude, server
Climate and Pace	<input type="text"/>	5 Key Words: leadership, product, focus, direct, senior, chang, leader, strategi, vision, silo, smart
Training and Support	<input type="text"/>	6 Key Words: hour, money, paid, vacat, overtime, holiday, call, student, mandatory, school, teacher, weekend
Growth Opportunities	<input type="text"/>	7 Key Words: learn, pace, need, experi, grow, differ, fast, profession, help
Work Shifts/Hours	<input type="text"/>	8 Key Words: environ, growth, advanc, easy, room, home, opportun, bore
Leadership Quality	<input type="text"/>	9 Key Words: manage, employe, care, upper, chang, treat, provid, listen, overwork
Vacation and Overtime	<input type="text"/>	10 Key Words: team, train, support, communic, organ, compens, improv, program, will

Note. Some of the words for the topic models are “**stemmed**”. This means that they have been reduced to their “base” or a simpler form. For example, “**cultur**” could be interpreted as “culture” and/or “cultural” and/or “cultured”; “**mobil**” as “mobile”, “mobilize”; “**vacat**” as “vacate”, “vacation”, “vacationing” and “**advanc**” as “advance”, “advancing”, “advancement”. Feel free to use your interpretation of the stemmed words to determine their intended meaning, but please reach out if you need further clarification on any of these stems if they prove to be unclear.

Write any additional thoughts/suggestions/observations here (optional):

Appendix B: Interaction Plots with Topic Proportions and Sentiment

Appendix B1. Work Shifts/Hours

Graph showing the relationship between the topic proportions for Work Shifts/Hours moderated by sentiment (N=24,539).

Appendix B2: Vacation and Overtime

Graph showing the relationship between the topic proportions for Vacation and Overtime moderated by sentiment (N=24,539).

Appendix B3: Company Culture/Climate

Graph showing the relationship between the topic proportions for Company Culture/Climate moderated by sentiment (N=24,539).