DISSERTATION

PROMOTING SAFETY THROUGH DIVERSITY MANAGEMENT: DIVERSITY CLIMATE, RACIAL ETHNICITY, AND SAFETY VOICE

Submitted by

Julia Lynn Otero Beckel

Department of Psychology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2024

Doctoral Committee:

Advisor: Danielle Gardner

Joshua Prasad
Mark Prince
John Rosecrance

ABSTRACT

PROMOTING SAFETY THROUGH DIVERSITY MANAGEMENT: DIVERSITY CLIMATE, RACIAL ETHNICITY, AND SAFETY VOICE

Given a sustained increase and undue burden of work-related injury and illness among racial-ethnic minorities in the United States, the current study aimed to provide a preliminary understanding into mechanisms which might lend to occupational health disparities. I utilized a two-wave survey approach via Prolific, a web-based survey platform, to gather perceptions from workers in high-risk industries (e.g., construction, manufacturing, healthcare, etc.) relating to their work groups' diversity climate, perceived organizational identification and psychological safety, worker safety voice, and occupational safety and health history. I considered four research questions: 1) How do employee perceptions of their organization's diversity climate relate to worker safety voice, 2) Do workers' perceptions of psychological safety and organizational commitment mediate the relationship between perceived diversity climate and worker safety voice, 3) What racial-ethnic differences exist in the indirect effects of diversity climate on worker safety voice via psychological safety and organizational identification?, and 4) If racial-ethnic difference in safety voice exist, are they also associated with differences in self-reported occurrences of accidents, injuries, or work-related illnesses across racial-ethnic groups? Results demonstrate the indirect effect of diversity climate on safety voice is significant via psychological safety, but not organizational identity. However, these effects did not influence occupational incident occurrence by race/ethnicity. I discuss relevant implications for theory and practice.

TABLE OF CONTENTS

PROMOTING SAFETY THROUGH DIVERSITY MANAGEMENT: DIVERSITY CLIMATE, RACIAL ETHINCITY, AND SAFETY VOICE

In the United States (U.S.), minority racial-ethnic groups hold an undue burden of work-related disability. In particular, Non-Hispanic Black workers of all ages are 1.3 times as likely to acquire a work-related disability than Non-Hispanic Whites (Seabury et al., 2017). Further, foreign-born Hispanic workers ages 18-29 are 2.2 times as likely, and those ages 50-64 are 1.7 times as likely, to develop work-related disability when compared to Non-Hispanic Whites. Yet, and despite the economic cost of workplace injury and illness in the U.S. (up to $250 billion a year; Seabury et al., 2017), work to better understand specific contributors to these differences is still sparse.

More broadly, racial-ethnic differences in work-related injury and illness might be explained through the occurrences of occupational health disparities. Occupational health disparities are "avoidable differences in work-related disease incidence, mental illness, or morbidity and mortality that are closely linked with social, economic, and/or environmental disadvantage such as work arrangements (e.g., contingent work), socio-demographic characteristics (e.g., age, sex, race, and class), and organizational factors (e.g., business size)" (National Institute for Occupational Safety and Health, 2019). Minority racial-ethnic groups are more commonly employed in high-risk occupations (e.g., manufacturing, food processing, healthcare). However, racial-ethnic differences in work-related injury and illness have been shown to occur even when minority groups work the same number of hours (Seabury et al., 2017), and in the same types of jobs (i.e., construction; Brown et al., 2021), as Non-Hispanic Whites.

Notwithstanding work done to exhibit the occurrence of occupational health disparities, less work has been done to identify the underlying mechanisms which contribute to racial-ethnic differences in injury and illness occurrence at the organizational-level (Landsbergis et al., 2014). Therefore, I propose to fill critical gaps in the literature by evaluating potential psycho-social contributors to racial-ethnic differences in workplace accident involvement and injury occurrence. Specifically, I intend to identify how worker perceptions of diversity climate lend to racial differences in worker safety voice – the act of speaking up about safety in order to prevent accidents and physical harm at work – via differences in perceived psychological safety and organizational identification. Further, I aim to evaluate whether potential differences in worker safety voice are also related to different levels of involvement in workplace accidents and injury occurrences among racial-ethnic groups. In doing so, I aim to contribute to the literature on occupational health disparities and safety voice in several ways.

First, the current study takes a novel approach to understanding the underlying mechanisms of safety voice and occupational health disparities by bridging the literature related to diversity in organizations and occupational health psychology. To date, there has been little work which has considered the individual and contextual differences that lend occupational health outcomes which may stem from whether workers feel comfortable voicing concerns or ideas related to safety at work. Work which has evaluated these processes has primarily been in relation to the experiences of young, versus adult workers, and has mainly helped to identify the role of supervisors in facilitating safety voice among workers with lessened power authority (i.e., young workers). In this study, I aim to further investigate antecedents and outcomes of safety voice expression by workers of various racial-ethnic identities in order to more directly address

occupational health disparities which occur across both young and adult working populations (Seabury et al., 2017).

Further, there are current gaps in understanding related to occupational health disparities and safety communication, specifically as it relates to safety voice. These gaps may be the result of several theoretical and methodological barriers. Methodologically, prior authors have questioned the appropriateness of using surveys and interview protocols when assessing safety voice due to the temporal nature of safety voice behaviors and the potential for common method bias to reduce the likelihood of developing a causal understanding of safety voice's antecedents and outcomes (Noort et at., 2019). While alternative methods such as observations may alleviate these concerns, it is also difficult to identify the outcomes of safety voice as remediation of safety hazards or implementation of safety procedures may be temporally distant from the voice behaviors that lead to them. To alleviate concerns of common method bias while also considering the proximal outcomes of worker safety voice, I use a multiple time-point design to gather perceptions and self-report behaviors from workers within high-risk occupations. Then, I evaluate the association between worker self-reported safety voice behaviors and their history of workplace injury, illness, or accident involvement while working for their current employer.

Finally, theoretical investigation into the mechanisms underlying occurrences of worker safety voice is nascent. Prior authors have begun to evaluate various antecedents to safety voice, as well as mediators of the relationship between safety voice and organizational outcomes, yet few authors have taken a theoretical approach to inform these processes. For instance, prior work has identified both psychological safety (O'Donovan et al., 2021) and organizational identification (Ni et al., 2022) as antecedents to the expression of safety voice. However, work to

date has yet to fully explain the process by which these affective states are manifested within employees and how they may contribute to safety voice expression.

While prior studies have referred to theories such as the job-demand resources model (Mathisen et al., 2022) to explain the facilitators and barriers to safety voice among the general workforce, no studies have addressed safety voice as an explanatory mechanism within an occupational health disparity framework. Subsequently, the literature lacks a theoretical perspective for addressing occupational health disparities more broadly. In the current study, I draw from Social Identity Theory (SIT) to attempt to mitigate this theoretical gap. Moving forward, I begin by expounding on occupational health disparities and related constructs among my current research effort. Next, I will discuss Social Identity Theory (SIT) as a relevant theoretical perspective to addressing worker safety voice and occupational health disparities, before outlining the current study's objective, hypotheses, and research design.

BACKGROUND AND SIGNIFICANCE

Occupational Health Disparities

Occupational health disparities reflect the inequitable proportion of work-related injuries, illnesses, and fatalities experienced by minority working populations (NIOSH, 2019), and in particular racial-ethnic minorities (Murray, 2003). Commonly, racial-ethnic minorities compose a disproportionate sample of workers among high-risk industries, such that workers of color are often underrepresented in "white-collar" professions and overrepresented among "blue-collar" and service industries. For example, in 2021, 42.6% of non-Hispanic White workers were employed in management, professional, and related occupations as compared to 33.9% of Black and 24.5% of Hispanic workers (Bureau of Labor Statistics, 2022). Meanwhile, a greater proportion of Black and Hispanic workers were also employed among service, production, transportation, and related industries which are often classified as "high-risk."

Accordingly, workforce participation rates among high-risk industries have been believed to explain the increased rate of adverse occupational outcomes among workers of minority racial/ethnic status. For instance, research comparing work-related injury and illness (WRII) claims from 2013-2017 in Washington state revealed a significant difference between WRII claim rates between non-White and White workers (Smith et al., 2022). In particular, the greatest claim-rate differences were documented among high-risk industries of agriculture, forestry, fishing, and construction. However, racial-ethnic minorities often demonstrate an undue burden of adverse occupational health outcomes above and beyond workforce participation rates (Seabury et al., 2017). In the same WRII investigation, both Black and Hispanic workers demonstrated increased claim rates across all industries (up to 14 times higher than non-Hispanic

White workers), regardless of the level of associated occupational risk, with Black workers demonstrating the highest rate of claims across all occupation groups.

Other statistical investigations also reflect an undue burden of occupational illness, injury and fatalities among racial-ethnic minorities. For example, 2020 data from the National Security Council (2022) reflects Hispanic and Latino workers experienced a death rate of 4.5 per 100,000 workers compared to 3.5 and 3.3 for non-Hispanic Black and non-Hispanic White workers, respectively. Similarly, based on results from the U.S. Census data, Seabury et al. (2017) report both Black and Hispanic workers have a significantly higher chance of developing a work-related disability as compared to non-Hispanic White workers, even when controlling for workforce participation rates. These disparities have been reflected through a number of other occupational-specific studies, with inequitable adverse occupational health outcomes occurring for racial-ethnic minorities in occupations from construction (Brown et al., 2016) to cleaning and hospitality (Premji & Krause, 2010). Given these findings, it is unlikely workforce participation is wholly accountable for explaining occupational health disparities in the United States, and thus researchers and practitioners are tasked with uncovering alternative mechanisms for their occurrence. In the current study, I evaluate the role of worker safety voice, or speaking up on safety at work, as an underlying factor contributing to the differential rates of occupational injury and illness among racial-ethnic minorities. Specifically, I investigate organizational factors which may explain why an individual chooses to speak up or remain silent regarding workplace safety, how these factors may differ based on one's race-ethnicity, and whether these differences help to expound on the differential rates of workplace injury, illness, and accident involvement among workers of minority race/ethnicity.

Safety Voice as an Explanatory Mechanism

Though uninvestigated to date, safety voice may serve as an explanatory mechanism for racial-ethnic differences in occupational injury and illnesses. Before discussing safety voice, it is helpful to understand the literature regarding employee voice more generally. *Employee voice* has been defined as "informal and discretionary communication by an employee of ideas, suggestions, concerns, information about problems, or opinions about work-related issues to person who might be able to take appropriate action, with the intent to bring about improvement or change" (Morrison, 2014, p. 174). Employee voice is assumed to be a form of extra-role performance via upward and constructive communication behaviors meant to change the status quo. The content of communication falls upon a spectrum, ranging from a presentation of ideas for improvement, referred to as promotive voice (Liang et al., 2012), to identifying problems within the workplace, termed prohibitive voice.

Distinct from employee voice, *safety voice* is an extra-role behavior which involves communication aimed at changing perceived unsafe working conditions (Noort et al. 2019). Safety voice occurs through both formal (e.g., filing an OSHA complaint) and informal (e.g., alerting a manager or coworker of hazardous work conditions) channels (Tucker er al., 2008), and may be directed towards varying levels of targets such as coworkers, supervisors and managers, and union officials. Additionally, the outcomes of an absence of safety voice can lend to severe outcomes for both worker and organizational health, including physical harm to both clients (e.g., patients within a healthcare setting) and workers (Noort et al., 2019). Conversely, increased levels of safety voice are associated with decreased safety risks and personal injuries (Mathisen et al., 2022).

Research suggests both individual and organizational characteristics can either enable or prevent a worker from speaking up on safety hazards in their work environment. For instance, in their review of the ecological nature of safety voice, Noort and colleagues (2019) discuss the role of employee demographics such as gender, job position, socioeconomic and power status, as well as workgroup culture and leadership styles for predicting the occurrence of safety voice in organizations. For example, occupations with traditionally dangerous working conditions are more likely to motivate workers to engage in safety voice behaviors (Hirschman, 1970), and demographically, women demonstrate higher levels of safety voice under working conditions with greater risks to health and safety (Tucker & Turner, 2014).

Similarly, Manapragada and Bruk-Lee (2016) discussed reasons employees may be likely to remain silent regarding safety issues, sometimes referred to as *safety silence*. They suggested workers are less likely to speak up on safety when there are, 1) fears of altering workplace relationships (e.g., retaliation), 2) perceptions of a weakened safety climate (i.e., perception the organization values safety and the behaviors, policies, and practices which demonstrate such value; Zohar, 1980), 3) issue-specific perceptions indicating a low likelihood of adverse outcomes, and 4) job-based factors such as having a high workload or a lack of trust in one's supervisor.

Theoretical Approaches to Safety Voice

More recently, research related to safety voice has begun to explore the psychosocial processes underlying safety voice via a variety of theoretical approaches. Tucker et al. (2014) drew upon social exchange approaches to safety to examine worker safety voice (Hoffman & Morgeson, 1999). In their study, the authors highlighted the importance of perceived organizational support for safety for encouraging worker safety voice behaviors. Specifically,

results from Tucker et al.'s study showed perceived coworker support for safety fully mediated the relationship between perceived organizational support and worker safety voice, such that one's coworkers serve as social exchange agents which represent the organization's commitment to safety, thus creating an environment in which attending to safety and demonstrating desirable safety behaviors becomes a social exchange.

Similarly, from a job-demands resources (Demerouti et al., 2001) perspective, Mathisen et al. (2022) hypothesized safety voice occurs as a result of decreased perceptions of job demands (i.e., workload, time pressure) and increased perceptions of job resources (i.e., supervisor support, job control). In their study, Mathisen et al. found job demands such as feeling overloaded with work tasks and perceptions of time pressure to be negatively related to safety voice expression, whereas positive job aspects, such as leader support and job control were positively associated with speaking up on safety. Notably, safety voice behaviors in their study were negatively associated with the outcome variables of safety risk (i.e., perceptions of unsafe work) and self-reported personal injuries.

However, while both Tucker et al. (2008) and Mathisen et al. (2022) take a theoretical approach to addressing the antecedents of safety voice and its outcomes, these studies still lack an explanatory approach to understanding why safety voice may vary based on individual differences such as racial-ethnicity. In the current study, I investigate precursors to safety voice using a diversity management approach incorporating potential organizational and psychosocial antecedents to safety voice, as well as any racial-ethnic differences which relate to the occurrence of worker safety voice.

The Potentiating Role of Diversity Climate

Given prior work demonstrating the social and psychological nature of safety voice, diversity climate may serve as unstudied mechanism for predicting the occurrence of safety voice behaviors as they relate to occupational health inequities between racial-ethnic minority and non-minority workers. *Diversity climate* refers to employees' perceptions that an organization values and supports diversity through fair employment practices and the extent to which minority employees are integrated within the work environment (Mor Barak et al. 1998). Among organizations with high levels of diversity climate perceptions (i.e., pro-diversity climates), you may expect the integration of various diversity management practices such as targeted recruitment methods aimed at gathering a diverse applicant pool, diversity trainings, or publicly displayed company diversity statements.

In the workplace, racial-ethnicity contributes to differing organizational experiences for employees of various identities (Singh et al., 2013). Accordingly, among racially and ethnically diverse workforces, diversity climate has been empirically linked to a number of positive performance behaviors among employees, including both increased task (McKay et al., 2007) and contextual (Singh et al., 2013) performance. When looking to distinguish the psychological mechanisms (i.e., "the black box") by which diversity climate influences employee performance, prior research has established the important role of both psychological safety (Sing et al., 2013) and organizational identification (Ashforth & Mael, 1989). However, these assumptions have yet to be extended to consider employee safety performance behaviors, such as safety voice. Yet, given the operational categorization of safety voice as a form of contextual performance, diversity climate and its associated benefits for employee behavior may serve as a relevant

antecedent to predicting differences in safety voice, and subsequently occupational health outcomes among racial-ethnic minorities.

THEORETICAL UNDERPINNINGS

Social Identity Theory

There is a growing belief that safety is socially constructed (Turnery & Gray, 2009) and involves a number of individual and interpersonal processes. Thus, Social Identity Theory (SIT) serves as a relevant theoretical and illustrative mechanism for understanding the social processes which might lend to racial-ethnic differences in safety voice behaviors resulting from diversity climate perceptions. Broadly, SIT posits individuals within a given social environment, such as one's workplace, rely on salient characteristics of others (i.e., race, gender, age) and the context of the environment to classify themselves within an in-group (i.e., "us") and out-group (i.e., "them") frame of reference (Tajfel & Turner, 1985). This classification process is meant to serve two primary purposes. First, social categorization provides individuals with an organizational lens for defining others within a given social environment, wherein individuals are likely to assign prototypical characteristics to others based on their perceived classification (i.e., stereotyping). Alternatively, social classification allows an individual the opportunity to define their own selves within the social environment. This definition constructs one's own social, rather than personal identity– or "the perception of oneness with or belonging to some human aggregate" (i.e., one's in-group; Ashforth & Mael, 1989, p. 21). The distinguishing feature of an individual's social identity then lies within the group or aggregate context (Tafjel & Turner, 1986). Wherein one's personal identity is largely predicated on interpersonal interactions where behavior is primarily under individual volition - one's social identity is primarily driven by group situations which allow for a subjective evaluation of intergroup comparisons.

In organizations, individuals may define themselves according to salient characteristics including one's age, gender, occupation, tenure, or race/ethnicity, as well as with the organization as its own entity. Accordingly, Ashforth and Mael (1989) provide an explanatory archetype of SIT as it relates to the organizational context, suggesting several factors which increase the likelihood of group, or organizational, identification. First, one vein of research suggests the distinctiveness of a group's values and practices in relation to comparable others facilitates the likelihood of an individual adopting the organization as part of one's greater identity (Mael, 1988; Ashforth & Mael, 1989). As an example, to engage in intergroup comparisons inherent to social identification, the distinctive features of an organization or workgroup should be conspicuous to the extent they allow for differentiation and define a clearly unique identity. Within an organizational context, hierarchical workgroups such as staff versus management might exhibit differential policies and practices which contribute to worker identification with their relative classification.

Another factor associated with social identification is the salience of relative out-groups, as easily distinguishable features of a given out-group reinforce individual awareness of one's in-group (Turner, 1981; Ashforth & Mael, 1989). Thus, awareness of a given out-group inherently demonstrates the existence of a boundary condition delineating "us" from "them", and subsequently strengthens the extent to which an individual identifies with their ingroup. In organizations, individuals may rely on salient features of the work environment such as the type of work (i.e., occupational identification) or observable organizational characteristics such as the demographic makeup of the organization's workforce (i.e., gender, race, age, etc.). For example, prior research by Kanter (1977) demonstrates how disproportional gender composition within the workforce strengthens ingroup identification. Specially, within a male-dominated workforce, the

presence of female workers was associated with greater perceived demonstration of masculine traits by male workers as well as greater perceived differences between female and male workers.

Finally, SIT posits a series of relevant outcomes associated with social identification. First, prior evidence suggests individuals are likely to choose activities and enact behaviors which align with their identity, as well as behaviors which support institutions embodying their identity (Mael, 1988). Further, social identification with a given group, or organization, is likely to increase outcomes associated with positive group formation such as feelings of altruism and a sense of loyalty to the group (Ashforth & Mael, 1989). Finally, SIT posits social identification should lend to individual acceptance of group values and norms, and increased homogeneity in attitudes and behaviors among a given group. Considering the current study, when presented with an organizational climate which supports employees' identities, individuals may be more likely to both socially identify with the organization, exhibit positive behaviors and attitudes toward the organization, as well as demonstrate support for the organization. For racial-ethnic minorities, in particular, this process may be represented through perceptions of a pro-diversity climate which demonstrates acceptance of and value for their personal identities, lending to positive affective and behavioral reactions.

RESEARCH QUESTIONS AND HYPOTHESIS DEVELOPMENT

Based on the above discussion, I address four research questions to establish an empirical link between diversity climate perceptions and worker safety voice, with implications for better understanding occupational health disparities more distinctly. In the following sections, I outline these questions and provide further empirical insight into my hypothesis development.

Diversity Climate as an Antecedent to Safety Voice

First, I investigate the following question: how do employee perceptions of their organization's diversity climate relate to worker safety voice (**R1)**? Following assumptions from Social Identity Theory (SIT), organizations demonstrating higher levels of pro-diversity climates (i.e., as indicated by participant ratings) should be more likely to better support employees' various social identities (i.e., racial-ethnic identity) and in turn encourage positive behavior and support toward the organization (Mael & Ashforth, 2002). Given the conceptual nature of safety voice as an extra-role performance behavior meant to direct positive safety-related change (Noort, 2019), workers perceiving pro-diversity climates in their organization should be more likely to engage in safety voice behaviors among their organization. Thus, I propose:

Hypothesis 1*:* Diversity climate perceptions will be positively and directly related to worker safety voice.

From Diversity Climate to Safety Voice

Next, I evaluate whether workers' perceptions of psychological safety and organizational identification mediate the relationship between perceived diversity climate and worker safety voice **(R2)**.

Psychological Safety

Psychological safety refers to the perception as to whether it is safe to take interpersonal risks (e.g., voice ideas, seek and provide feedback; Edmondson, 1999), and has been shown to predict the extent to which employees engage in promotive and prohibitive voice behaviors within workgroups (Bienefeld & Grote, 2014; Liang et al., 2012). When workers and employees have a greater sense of psychological safety, they are more likely to engage in voice behaviors overall (Edmondson, 2004). On the other hand, when employees feel unsafe taking interpersonal risks at work, the likelihood of employees engaging in either promotive or prohibitive voice behaviors decreases. For instance, authors have proposed workers facing job insecurity or socioeconomic disparity may be less likely to speak up on safety issues, or stop work upon discovering safety hazards, due to fear of retribution surrounding job loss (Landsbergis et al., 2014). Accordingly, involuntary part-time workers who frequently experience job insecurity, have demonstrated low levels of employee voice potentially due to concerns of evoking negative reactions from management (Stamper & VanDyne, 2007).

Given assertions from social identity researchers, participants reporting more positive ratings of perceived diversity climate should also report higher ratings of psychological safety due to the identity affirming effects of pro-diversity climates for all employees (Ashforth & Mael, 2002). For example, in organizations in which minority groups are degraded, White-identifying employees also report discomfort within the organization (Buttner et al., 2009). Additionally, when diversity is de-valued within organizations, employees are likely to restrict self-expression and interpersonal risk-taking (Foley et al., 2002). On the other hand, in organizations which value diversity and provide opportunities for positive social interactions, workers are more likely to perceive reduced threat, experience increased interpersonal liking, and

subsequently engage in individual categorization into social groups (i.e., one's work group; Asforth & Mael, 1989; Ashforth et al., 2008). Self-categorization and identification with one's workgroup should lend to increased psychological safety, and subsequently encourage extra-role behavior and the likelihood to take interpersonal risks related to voicing safety concerns at work.

Psychological Safety and Safety Voice. Like employee voice, researchers have recently begun to posit psychological safety as a predictor to safety voice. For example, although safety voice is an important factor for avoiding accidents in the workplace, research indicates employees will often choose to remain silent regarding safety hazards due to the potential unwanted consequences of safety voice (Liu et al., 2010). Nonetheless, recent research demonstrates psychological safety may lie amongst the root of the issue reduced safety voice. For instance, results from Tucker and Turner's (2014) study among young workers demonstrate psychological safety moderated the relationship between hazardous working conditions and the likelihood of young workers engaging in safety voice, such that those within hazardous working environments were less likely to engage in safety voice when psychological safety was low. Conversely, Sun et al. (2022) reported psychological safety to positively mediate the relationship between organizational climate and safety voice, such that high levels of organizational safety climate were related to greater reports of psychological safety, and subsequently greater intentions to engage in safety voice behaviors. Given these findings, there lies the potential for psychological safety to play an explanatory role in the proposed relationship between diversity climate and safety voice, such that positive organizational perceptions may more likely influence individual perceptions of psychological safety, resulting in a greater likelihood to engage in safety voice behaviors.

Given these findings, I investigate whether the indirect effect of diversity climate via psychological safety holds when evaluating the occurrence of worker safety voice behaviors. Considering past findings and the conceptual similarities, though distinct taxonomies, between safety voice and general employee voice, as well as the conceptualization of safety voice as an extra-role behavior, pro-diversity climates might also influence the extent to which workers speak up on safety at work via higher levels perceived psychological safety.

Organizational Identification

Relatedly, organizational identification refers to a psychological state representing the latent bond between an individual and the organization (Edwards, 2005), and is operationalized as a more specific form of social identification (Ashforth & Mael, 1989). In other words, organizational identification reflects a worker's "perception of oneness with, or belongingness to the organization" (p. 22). Among the organizational behavior and psychological literature, organization identification has been shown to be meta-analytically and positively associated with the likelihood of employee voice behaviors (Ruan & Chen, 2021). For example, both organizational identification as well as similar constructs of social and personal identification have been shown to mediate the relationship between transformational leadership and employee voice behaviors (Liu et al., 2010; Hu et al., 2015), as well as ethical leadership and voice behavior (Islam et al., 2018). Organizational identification has also been shown to play as a moderator in the relationship between a number of work-related factors and employee voice, including psychological safety (Wang et al., 2018) and job control (Tangirala & Ramanujam, 2008). Across these studies, therein lies the assumption that psychological processes which increase positive perceptions of the organization also influence the extent to which an individual

identifies with the organization (Mael & Ashforth, 1989), subsequently increasing the likelihood of employees to engage in contextual performance behaviors such as employee voice.

Organizational Identification and Safety Voice. Regarding safety voice, research investigating the role of organizational identification is cursory. For instance, Hu and Casey (2016) reported organizational identification to be associated, though not directly predictive of employee safety voice. However, Hu and Casey's study utilized a cross-sectional design and relatively small sample size (N = 165), suggesting the potential for common-method bias (Podsakoff & Olof, 1986) and insufficient power to detect an estimated effect size. Given this is the only study to date to directly evaluate the relationship between organizational identification and safety voice, more research with advanced research designs is needed to rule out a predictive relationship. In the current study, I aim to evaluate the mediating role of organizational identification in the relationship between diversity climate and safety voice, such that diversity climate should lend to increased levels of organizational identification, and subsequently increased reports of safety voice, given the conceptual similarity between safety voice and employee voice, broadly.

Considering the preceding discussions related to psychological safety and organizational identification, I hypothesize:

Hypothesis 2: Diversity climate will be positively and indirectly related to worker safety voice via (a) psychological safety and (b) organizational identification, such that workers reporting pro-diversity climates will also report higher levels of these constructs, and subsequently higher levels of safety voice.

Racial Ethnic Differences in Safety Voice

Next, I examine what racial-ethnic differences might exist in the indirect effects of diversity climate on worker safety voice via psychological safety and organizational identification **(R3)**.

Diversity Climate, Psychological Safety, and Racial/Ethnic Identity

For racial-ethnic minorities in particular, psychological safety plays an explanatory role in the relationship between diversity climate and the occurrence of extra- and in-role performance behaviors above and beyond that of non-Hispanic Whites (Singh et al. 2013; McKay et al., 2008). Specifically, greater perceptions of diversity climate have been indirectly associated with workers' organizational citizenship behaviors via psychological safety for workers of all racial-ethnic backgrounds, and this indirect effect is greater among workers of minority racial-ethnic identities (Singh et al., 2013).

Further, prior research suggests that employees within racial-ethnic minority groups are more likely to experience racial discrimination and to also identify more strongly with their race than White individuals (Clark et al., 1999). On the other hand, workers of majority status (i.e., non-Hispanic White employees), are less likely to experience racial discrimination and therefore may hold less value for pro-diversity work environments (Merritt et al., 2010). Accordingly, research suggests racial minorities more strongly identify with their organizations when perceiving a strong as compared to low diversity climate (Cole et al., 2016), and respond more strongly and positively to pro-diversity climates than racial majority workers because they affirm their social identity which makes them feel psychologically safe, and subsequently more likely to engage in behaviors which support the organization (De Cremer, 2002; Ely & Thomas, 2001; McKay et al., 2008).

Thus, within the context of this study, racial-ethnic minorities working within a pro-diversity climate should perceive higher levels of psychological safety and experience greater organizational identification, and subsequently report greater levels of safety voice within their workgroup than non-Hispanic White workers. Further, both Hispanic as well as non-Hispanic Black workers have historically demonstrated the greatest rates, controlling for workforce participation, of occupational injury and illness in the United States.

Diversity Climate, Organizational Identification, and Racial/Ethnic Identity

Considering racial-ethnic minorities, organizational identification has been shown to play a particularly important association in several work-related phenomena. In particular, Cole et al. (2016) report diversity climate to positively predict organizational identification, moderated by race-ethnicity, such that the relationship between diversity climate and organizational identification was stronger for racial/ethnic minorities than non-minority workers. These results mirror other studies in which minority employees demonstrate greater value for pro-diversity climates, such that racial-ethnic minorities may be more attune to organizational signals, policies, and practices relating to diversity (McKay et al., 2007).

Other studies have documented an association between organizational identification and diversity climate, such that diversity climate moderates the relationship between demographic and categorical (i.e., race and gender) demography and organizational identification (Gonzalez & Denisi, 2009); Specifically in Gonzalez and Denisi's study, Black participants reported higher levels of organizational identification when also reporting high perceptions of diversity climate; this effect was non-significant for White participants. Other research demonstrates pro-diversity climates are positively associated with organizational identification among minority group members within culturally diverse organizations (Hofhuis et al., 2012). Given the empirical

linkage between diversity climate and organizational identification, in the current study I look to expand this phenomenon by investigating the mediating role of organizational identification in the relationship between diversity climate and safety voice.

Given these discussions, I hypothesize:

Hypothesis 3: The proposed direct and indirect effects of diversity climate on safety voice will be the strongest for participants holding a racial-ethnic minority status, with non-Hispanic Black participants demonstrating the strongest effects, followed by both foreign- and native-born Hispanic, and non-Hispanic White participants, respectively.

Tying the Knot on Occupational Health Disparities

Finally, in order to demonstrate an empirical linkage between racial-ethnic differences in safety voice lend to the occurrence of occupational health disparities, I investigate if racial-ethnic differences in safety voice exist, are they also associated with differences in self-reported occurrences of accidents, injuries, or work-related illnesses across racial-ethnic groups (**R4)**? Given the long-standing documentation of safety voice as an antecedent to occupational injuries (Noort et al., 2019), I propose:

Hypothesis 4: Demonstrable effects in the indirect effect of diversity climate on safety voice via organizational identification and psychological safety will also lend to differences in reported occupational incidents by racial ethnicity.

I present the proposed relationships for all hypotheses within Figure 1. In the following section, I describe steps for evaluating these proposed hypotheses using a multi-time point survey, encompassing minority and majority racial/ethnic groups across high-risk occupations.

METHOD

Participants

Recruiting

I recruited participants and gathered data within Prolific, a web-based crowdsourcing platform. Web-based platforms such as Prolific allow researchers to recruit a wide and demographically diverse participant sample and can greatly reduce recruiting efforts for achieving a diverse sample through controlling and screening for a number of factors including race/ethnicity and occupational sector. Additionally, Prolific has been shown to demonstrate high retention rates among longitudinal samples (approximately 90%, Peer et al., 2017), resulting in high-quality data.

Eligibility criteria was set such that participants were to be of adult age, identify as Non-Hispanic White, Non-Hispanic Black, or Hispanic, and report working within a high-risk industry as indicated by the Bureau of Labor Statistics (e.g., construction, manufacturing, food processing, healthcare, etc.) as well as an occupation requiring the use of safety protocols. Participants were recruited using quota sampling procedures, with a maximum sub-group sample size of n = 200 by race-ethnicity based on an a-prior power analysis suggesting a total sample size requirement of N = 550 (B = .80) for detecting prespecified standardized path estimates at *effect* = .16, the median effect size among pre-registered psychological studies (Schafer & Schwarz, 2019). However due to attrition across timepoints, my final sample was just below this threshold.

The original T1 sample was comprised of N = 560 respondents (Non-Hispanic White: n = 200, 26%; Non-Hispanic Black: n = 200, 36%; Hispanic n = 160, 28%). Of those respondents, 67

respondents did not return for the second survey, with Hispanic participants showing the greatest attrition rate (Non-Hispanic White: n = 6; Non-Hispanic Black: n = 24; Hispanic: n = 23). Thus, the final sample included 493 working adults in the U.S. who identified as non-Hispanic White (39.6%), Non-Hispanic Black (35.7%), and native or foreign-born Hispanic (24.7%). Participants were on-average, 38.6 years old (*SD* = 11.1), primarily identified as Male (60.0%), and most commonly worked in Health Services (27.18%), Manufacturing (13.39%), Retail Trade (12.79%), and Construction (11.97%). I present additional participant demographics in Table 1, and a breakdown of demographics by Race/Ethnicity in Tables 2-4.

Procedure

Design

I implemented a series of online questionnaires using a two-wave survey design, with a one-month time interval between surveys. The decision to incorporate multiple time-points for data collections is largely due to circumstances in which post-hoc reports of psychosocial phenomenon are subject to memory bias and common method variance (Podsakoff & Oloff, 1986). *Common method variance* refers to the problems which arise when collecting multiple measures of psychological or post-hoc phenomena from the same participant. In instances in which measures are collected at the same time, from the same source (i.e., participant), any deficiencies in that source are assumed to contaminate each measure, most likely to the same degree and direction. Thus, by collecting exogenous and endogenous variables across multiple time points, I am able to detect and control for common method variance resulting from spurious influences such as participant mood states or stimulus cues resulting from the ordering of included measures (Podsakoff et al., 2012).

Finally, assessing variables across time-points allows me to assess temporal, versus atemporal mediation. The primary assumption under temporal mediation is that the relationship between variables is irreversible and sequential, allowing for causal interpretations (Winer et al., 2016). Under this assumption, Variable X caused Variable M, which in turn caused Variable Y, chronologically. Thus, by utilizing scale scores relating to the variables of interests across separate time points, I might draw more generalizable and causal interpretations of the proposed study results than if using a cross-sectional approach.

Data Collection

Within Prolific, participants were pre-screened according to the eligibility criteria. Those participants meeting the eligibility criteria were able to access the online informed consent materials and first timepoint survey. This survey included items related to demographic and occupational characteristics, as well as scales corresponding to diversity climate, organizational identification, psychological safety, and safety voice. One month later, participants were invited to return to complete a second survey, which included all prior scales as well as items assessing their occupational health and safety history within the one-month time lag between assessments. I compensated participants $2.00 per fully completed survey, including a correct response to at least one of two attention checks (notably, all participants met attention check criteria). This payment rate corresponded to an expected 10-minute completion time at the United States' average minimum wage of $12.00/hr. Finally, with intent to bolster participant retention across surveys, participants who completed both surveys received an additional $1.00.

Measures

Participants responded to measures corresponding to the variables of interest (i.e.,

diversity climate, psychological safety, safety voice behaviors, demographics, occupational health and safety history) as well as measures relating to characteristics of their job and organization. Items for each measure are reported in Appendix A.

Diversity Climate

I measured diversity climate via a nine-item scale adapted from McKay et al. (2007). Participants were instructed to indicate the extent to which their organization meets their own diversity-related expectations and respond via a five-point Likert scale. Possible responses ranged from (1) *Well below expectations* to (5) *Well above expectations*. In McKay et al. (2007) study, the scale demonstrated excellent internal consistency ($\alpha = .91$), as well as in the current study ($\alpha = .93$).

Organizational Identification

To measure participant perceptions of organizational identification, I adapted Mael and Ashforth's (1992) six-item organizational identification scale to be used within a work-specific context. Participants responded on a scale from (1) *Strongly disagree* to (5) *Strongly agree*. In the current study, the measure demonstrated exceptional internal consistency ($\alpha = .91$).

Psychological Safety

To gather psychological safety perceptions at the individual-level, I utilize an adapted version of the six-item team-level psychological safety scale by Edmondson (1999). Adapting this team-level scale is a common practice among researchers interested measuring psychological safety among individual workers (Edmonson, 2004). Response options ranged from (1) *Strongly disagree* to (5) *Strongly agree*. In the current study, the measure demonstrated acceptable reliability ($\alpha = .72$).

Safety Voice

While conceptualization of the construct of safety voice is relatively nascent, so is the field's understanding of best practices for measuring safety voice behaviors. Researchers have pursued a number of qualitative and quantitative approaches to measuring safety voice in organizations. To date, post-hoc and self-report methods have been determined to be useful for circumventing ethical concerns regarding lab-based safety voice studies and are the primary mechanism for measuring safety voice in organizations (Noort, 2019). I utilized an adapted version of the six-item Safety Voice scale developed by Tucker et al. (2008), which has been shown to demonstrate sufficient internal consistency ($\alpha = .78$) in prior and the current study ($\alpha = .87$). Response options ranged from (1) *Strongly disagree* to (5) *Strongly agree*.

Occupational Safety Incident Involvement

To investigate the relationship between safety voice and actual occurrences of occupational injury/illness, participants were asked to report the number of work-related injuries or illnesses they acquired on the job while working for their current organization. I adapted items from the National Institute of Occupational Safety and Health's WellBQ (Chari et al., 2021). Examples include slips, trips, and falls of all severity, repetitive movement injuries, Covid-19 when contracted from one's worksite, heat exhaustion/stroke, among others. Frequency of reported incidents ranged from 0-5. Participants were also asked to rate the severity of their worst workplace illness or injury on a Likert scale ranging from (1) *Not at all severe* to (5) *Very severe*, as well as qualitatively explain the context and type of reported illness/injury. Finally, participants were asked to report the number of near miss incidents, or close calls, in which they

were involved during the one-month between their first and second surveys. Frequency of reported near misses ranged from 0-5.

Racial-Ethnicity

In order to sufficiently capture the demographic samples reflected among my hypothesis, I collected participant self-reported race and ethnicity. Prolific eligibility criteria pre-screened participants based on the race/ethnicity indicated within their user-profile. Those meeting the criteria for inclusion (i.e., those identifying as non-Hispanic White, non-Hispanic Black, native or foreign-born Hispanic), were asked to report their race and ethnicity a second time in order to verify eligibility.

Control Variables

Finally, I collected participant demographics and organizational characteristics as potential control variables. Specific variables included age, gender, tenure, and self-report perceptions of their organization's safety climate ($\alpha = .86$ in the current study) due to their predictive role related to safety voice behaviors reported in prior literatures (Noort, 2019).

Analysis Plan

In the following sections, I describe the analytical steps taken to test my hypotheses. I first explain steps taken to evaluate psychometric characteristics via confirmatory factor analysis and invariance testing across groups and time. Next, I describe how I use an extension of a half-longitudinal panel design (Kline, 2015; Cole & Maxwell, 2003) for hypothesis testing and include a discussion of why a half-longitudinal design is an appropriate analytical method given the structure and context of the data.

Confirmatory Factor Analysis

I began by conducting a series of confirmatory factor analyses (CFA) to assess the factor structure of each latent variable. CFA is a type of factor analysis that demonstrates the relationship between observed indicators (i.e., scale items) and the latent constructs or factors they are proposed to measure (Nye, 2023). This method helps to determine whether the latent factors associated with a set of indicators align with the theoretical conceptualization of the construct. As such, CFA is most appropriate when there is a well-known a priori factor structure with established constructs. There are a variety of other factor analysis methods that each serve a unique purpose. For example, exploratory factor analyses (EFA) may be used when the a priori factor structure of a construct is unknown and is most useful for constructs that have limited research support (Nye, 2023). Rather in this study, where each construct was measured with previously validated scales, CFA is the preferred choice of analysis.

I conducted a separate CFA for each latent study variable, at each time point. Each variable was modeled to represent a single factor except for *safety voice* (Tucker et al., 2008), which was specified to represent two underlying factors: three items representative of *promotive* voice, and three items representing *prohibitive* voice. For each latent factor, I fixed the first factor loading to 1. Since the latent factors in a specified model lack a known metric, researchers must assign one to estimate the model (Nye, 2023). By fixing the first factor loading to 1, the scale of the latent variable is then set to be the same as the indicator with the fixed loading.

To evaluate model fit, I referred to fit indices criteria suggested by Hu and Bentler (1999). These indices include the comparative fit index (CFI) > .95, Tucker-Lewis Index (TLI) > .95, root mean square error of approximation (RMSEA) < .06, and standardized root mean square residual (SRMR) < .05. Based on best practice reporting methods by Jackson et al.

(2009), model fit was deemed adequate when at least one measure of incremental fit (e.g., TLI, CFI) and one measure of residualized fit (e.g., RMSEA, SRMR) met the fit indices criteria.

In addition, I report the Chi-Square test of model fit, where a non-significant test indicates perfect fit, or no difference between the proposed factor model and the observed data (Barrett, 2006). However, past findings suggest that large sample sizes frequently lead to Chi-Square tests signaling model misfit unnecessarily (Nye, 2023). Instead of using the results of the Chi-Square test as a binary fit/no-fit decision rule (Barrett, 2006), I present the results of the Chi-Square test along with the fit indices criteria by Hu & Bentler (1999) for transparency.

Invariance Testing Across Groups and Time

Measurement equivalence across groups. Next, I evaluated measurement equivalence via a series of invariance tests following recommendations by Kline (2011). This procedure involves a CFA approach wherein increasingly stringent constraints are added at each level of CFA testing, including configural tests (equivalence of structure), metric tests (equivalence of factor loadings), and scalar tests (equivalence of item intercepts; Vandenberg & Lance, 2000). For each latent variable, I tested whether the factor models tested above were equivalent between White, Black, and Hispanic racial/ethnic groups on the configural, metric, and scalar levels.

Specifically, for each latent variable I began by performing a series of CFAs by racial-ethnicity to evaluate and ensure adequate model fit by group (i.e., White, Black, Hispanic). Since data collection occurred within Prolific, where participants were pre-screened by racial/ethnicity, all participants were categorized into one of these three groups. Next, I tested for configural invariance by performing a multi-group CFA with all parameters allowed to vary (i.e., no equality constraints; Vandenberg & Lance, 2000). To determine whether the constructs were equivalent at the configural level, I first evaluated model fit based on the fit indices described in

the prior section (Hu & Bentler, 199) and the Chi-Square goodness-of-fit test. Generally, if adequate model fit is not reached such that the factor structure significantly varies across groups, researchers are advised to stop testing as these results imply the constructs are fundamentally different across groups (Vandenberg & Lance, 2000).

When adequate model fit is achieved, the next step is to test for metric equivalence by re-specifying the model and constraining the factor loadings to be equal across groups (Vandenberg & Lance, 2000). When applying additional constraints to the model, model fit will diminish. However, achieving metric equivalence can be determined by 1) adequate model fit and 2) non-significant changes in model fit as determined by Chi-Square goodness-of-fit testing. In the instance the difference in goodness-of-fit between levels is non-significant, researchers may begin to speak to the level of invariance achieved and subsequently continue to apply increasingly strict constraints (i.e., equivalence of item intercepts) until model fit is judged as inadequate and there is a significant change in the adequacy of model fit as determined by the Chi-Square goodness-of-fit test. Under these circumstances, achieving measurement equivalence provides varying levels (i.e., configural, metric, scalar) of evidence that the differences in scores across groups are due to meaningful differences in the variable measured rather than differences in the measurement's performance.

Measurement equivalence across time. I also tested measurement equivalence across time points for each latent variable (i.e., *longitudinal measurement equivalence;* Little et al., 2007*).* In longitudinal research, measurement equivalence across time indicates changes in scores across time points are due to actual changes in the construct being measured, rather than changes in the measurement over time. Simply put, the underlying question is whether the indicators represent the same latent factors over time. As with between-groups measurement

equivalence, longitudinal measurement equivalence involves a similar series of tests to examine equivalence (i.e., configural, metric, scalar). To note, these tests evaluate equivalence across timepoints using the entire sample rather than equivalence within a given timepoint, but across groups.

Like the CFA procedures used for testing measurement equivalence across groups, I began by specifying an unconstrained model (i.e., no equality constraints) to test for configural invariance. However, rather than setting my grouping variable to separate groups by racial/ethnicity, I specified separation of observations by timepoint (i.e., time 1, time 2). I also followed similar procedures to assess whether the constructs were equivalent at the configural level across timepoints by referring to best-practice fit indices (Hu & Bentler, 1999) and the Chi-Square goodness-of-fit-test (Little et al., 2007). When testing for longitudinal invariance, researchers are advised to at a minimum meet requirement for metric invariance (i.e., equality of loadings) before hypothesis testing (Little et al., 2007).

Following these steps, I followed the invariance testing procedures outlined in the prior section, by testing for increasingly strict equality of constraints (i.e., equality of loadings, equality of intercepts) and evaluating model fit using the same criteria (Hu & Bentler, 1999; Little et al., 2007). As discussed, adding equality constraints will naturally depreciate model fit, so for each subsequent model equivalence was determined via adequate model fit and a non-significant change in model fit via the Chi-Square goodness-of-fit test. Upon determining inadequate model fit or a significant change in model fit due to the addition of equality constraints, I ceased testing stricter models.

Hypothesis Testing

To test my hypotheses, I used structural equation modeling via an extension of a half-longitudinal design (Cole & Maxwell, 2003). In a full-longitudinal design, predictor (X), mediator (M), and outcome (Y) variables are measured on a minimum of three different occasions (Cole & Maxwell, 2003). In traditional half-longitudinal designs, M and Y are measured at both time points, whereas X is measured only at Time 1 (Kline, 2015). Alternatively, X and M may be measured contemporaneously, and Y is measured only once. Therefore, full-longitudinal models can be framed as the combined outcome of two half-longitudinal models. For the current study, I utilized a half-longitudinal design with two extensions: a first stage latent moderation effect and serial mediation extending the b-path to test the path of safety voice on incident occurrence.

Considerations in half-longitudinal designs. There are several factors to be considered when using a half-longitudinal design for assessing mediation. Collins et al. (1998) posit that three waves are necessary to test mediation and provide time precedence. This assumption is rational when measuring stage-based or transitional phenomena common to randomized treatment designs. With stage-based phenomena, the causal process is assumed to begin when X is measured and requires additional measurement points to assess the effect of X on M, and M on Y. However, Kline (2015) states time precedence is not guaranteed. The covariance between X and Y can be large even if Y causes X and the effect (X) is measured before the cause (Y). Kline argues this phenomenon is possible because X can still be affected by Y before either variable is measured in a longitudinal study.

In addition, Cole and Maxwell (2003) argue that half-longitudinal designs are appropriate when evaluating phenomena that represent an ongoing process. For example, the current study

assumes the relationships between participant perceptions of diversity climate (X), psychological safety (M), organizational identity (M), and safety voice (Y) are ongoing and developed before the first assessment. Cole and Maxwell describe this concept as *stationarity,* the "constancy of means, causal processes, and variances-covariances over time" (p. 560). When researchers determine a phenomenon conceptually meets the stationarity assumptions, they can specify paired longitudinal tests to overcome the bias that Collins et al. (1998) assume characterizes the half-longitudinal design: 1) a-path estimation of M2 (i.e., time 2) onto X1 controlling for M1 (i.e., time 1), and 2) b-path estimation of Y2 onto M1 controlling for Y1. If the assumption of stationarity is met, the b-path between M1 and Y2 would then be equal to the b-path between M2 and Y3, and the product *a*b* provides an estimate of the mediation effect of X on Y through M. At times, these paired longitudinal tests are referred to as an *autoregressive model* (Mackinnon et al., 2007).

In the current study, these paired longitudinal tests are represented through the following specifications: 1) a-path estimate of time 2 (T2) mediating variables (i.e., organizational commitment and psychological safety) onto time 1 (T1) predictor (i.e., diversity climate), controlling for T1 levels of the mediators, and 2) b-path estimation of the T2 outcome variable (i.e., safety voice) onto the T1 measures of both mediating variables (i.e., organizational commitment and psychological safety), controlling for T1 measure of the outcome variable.

Two final considerations emerge when using half-longitudinal designs (Cole & Maxwell, 2003). First, using the specifications above we can only test whether M is a partial, rather than complete, mediator of the X-Y relationship. Second, if the stationarity assumption is not met then the relative path estimates will be biased. However, without three waves of data, such as in the current two-wave design, we cannot directly test the stationarity assumption. Nonetheless,

authors agree that failing to consider prior levels of the dependent variable will create more bias than violating the stationarity assumption, and a half-longitudinal model provides a superior test of mediation than cross-sectional design. As stated by Cole and Maxwell (2003) "…failing to control for prior levels of the dependent variables typically creates much greater problems than does failing to take into account violations of stationarity" (p. 563).

Alternative approaches. There are alternative approaches to assessing mediation with longitudinal data, specifically data collected via two-wave designs. One model is the latent-growth modeling (LGM) or parallel-process model (Mackinnon et al., 2007). LGM is used to evaluate whether the growth in predictor (X) affects growth seen in the mediator (M) which effects growth of the outcome variable (Y). However, the mediational relationship under these models is seen as correlational, such that that growth (i.e., change) in M is interpreted as correlated to the growth in Y at the same time (i.e., parallel process). The disadvantage with this model over the half-longitudinal auto-regressive model is an inability to assess and control for the influence of M at earlier levels or demonstrate the longitudinal effect of M on Y across timepoints. To mitigate this factor, researchers may specify a *piecewise parallel process model* (Cheong et al., 2003), where the relationship between the growth in M and change in Y are modeled for earlier and later times separately. Yet, this model also disallows the ability to concurrently evaluate the longitudinal relationship of M and Y across timepoints.

Another approach, the *latent difference score (LDS)* model specifies the difference scores between X, M, and Y and then analyzed using cross-sectional techniques. The assumption behind LDS models is that the latent difference scores are seen as differences between waves of

change. A benefit of the LDS model is its usefulness when predictors vary across measurement occasions. However, the model may still permit bias associated with cross-sectional techniques.

Thus, in consideration of the proposed research questions, the structure of the data collected (i.e., two-wave data), and the alternative analytical options for hypothesis testing (i.e., cross-sectional models) the half-longitudinal design is the most appropriate analytical procedure.

Model specification. To test H1, I specified diversity climate [Time 1 (T1)] as predicting safety voice [Time 2 (T2)]. For H2, I specified diversity climate (T1) as predicting psychological safety (T2) and organizational identification (T2). Subsequently, psychological safety (T1) and organizational identification (T1) were both specified to predict safety voice. (T2). To test H3, I began by creating the product term between my predictor (i.e., diversity climate) and two dummy-coded variables representing my moderator (i.e., non-Hispanic Black, Hispanic)[1]. Then, I specified these product terms as predicting psychological safety (T2) and organizational identification (T2). Finally, to test H4 I specified safety voice (T1) as predicting the number of reported incidents (T2). Finally, to control for prior levels of each predictor variable, I specified autoregressive paths of each T1 variable as predicting itself at T2 (i.e., psychological safety T1 predicting psychological safety T2). Doing this allows for the specification of a cross-lagged autoregressive model (i.e., separation of time-points) to mitigate concerns of common-method bias (Podsakoff et al. 2012), as well as provide an estimate of prior levels of each predictor.

[1] Latent Moderated Structural Equation Model (LMS) procedures were used to estimate interaction effects. Compared to regression-based methods, LMS procedures are considered advantageous such that they allow for the incorporation of measurement error and offer the greatest power to detect effects (Sardeshmukh & Vandenberg, 2017).

Model fit with Bayesian estimation. When estimating my model, I chose to specify a Bayesian estimator rather than alternative SEM methods which apply maximum likelihood (ML) estimation. Modern authors argue that analyses using ML and likelihood-ratio χ^2 testing apply unnecessarily strict constraints which lead to rejection of the model (Muthen & Asparouhov, 2012). Subsequently, authors may then rely on model modifications which often rely on chance. Further, when specifying indirect effects, ML provides parameter estimates and standard errors based on large-sample theory which assumes the distribution of a parameter estimate is normal (MacKinnon, 2007). Rather, using a Bayes estimator provides the whole distribution, or the posterior distribution, without assumption of normality (Muthén & Asparouhov, 2012). This allows for a credibility interval which may facilitate a strongly skewed distribution, rather than the ML confidence interval which assumes symmetry.

When using Mplus 8.4, Markov Chain Monte Carlo simulations allows for the exploration of the posterior distribution of Bayesian models via Potential Scale Reduction (PSR) criteria. Specifically, MCHC estimates posterior values via a series of iterations drawn from a prior distribution (Zitmann & Hecht, 2019). The PSR allows researchers to assess convergence via an estimate of precision (Gelmin & Rubin, 1992). In Mplus, 8.4, the default PSR criterion is PSR < 1.10. To inspect the PSR, Mplus procedures require a second model estimation with an increased number of MCHC iterations. Precise convergence is then determined by whether the PSR value does not increase and remains within the recommended range (PSR < 1.10) (Zitmann & Hecht, 2019).

RESULTS

Descriptive Statistics

Prior to hypothesis testing, I began checking for item, scale, and person-level missingness (Newman, 2014). I did not identify any missing data for key study variables. Additionally, I inspected outlying scale means of (+-) 2 standard deviations and removed these cases from future analyses (i.e., 24 cases). All variables were normally distributed as each met cutoffs for skew (-2 > X < +2) and kurtosis (-3 > X < +3), except for reported incidents occurred (Kline, 2011; Yuan & Bentler, 2000). This item experienced substantial variability given the low number of respondents reporting having experienced an incident within the study timeframe (i.e., 78 participants).

After removing outliers, the final sample for analysis was N = 469. I report descriptive statistics in Table 5, and correlations for all variables, including proposed control variables, in Table 6. As depicted in Table 6, *safety climate* had moderate to strong, positive correlations with multiple predictors and outcomes (i.e., diversity climate, psychological safety, organizational identification, and safety voice). These relationships were in the expected direction. Two demographic characteristics, *age* and *tenure* also shared small correlations with key study variables (see Table 6), but neither demonstrated relationships with *diversity climate.*

Notably, identifying as a racial-ethnic minority (i.e., Non-Hispanic Black and/or Hispanic), was significantly and positively correlated with the number of occupational incidents (r = .107, p = .022) participants experienced. Identifying as Non-Hispanic Black as compared to White was also significantly related to the number of incidents occurred (r = .109, p = .020). Although identifying as Hispanic compared to White was not significantly correlated with

occupational incidents reported (r = .082, p = .162). Total and average number of incidents and near misses by race/ethnicity can be found in Table 7.

Latent Variable Modeling

Confirmatory Factor Analysis

The measurement model for each latent variable were estimated using the full study sample (i.e., all racial-ethnic groups). All scales exhibited sufficient reliability (alpha >.70). Model fit indices, and results from Chi-Square goodness-of-fit testing are presented for all variables and timepoints in Table 8.

Diversity climate. The measurement model specifying a single factor of *diversity climate* from the first timepoint (T1) resulted in mediocre fit (RMSEA = .161 [.148, .174], CFI = .842, TLI = .791, SRMR = .064). The Chi-Square test of model fit was significant (χ^2 (27) = 346.304, p = .000). The standardized factor loadings ranged from .680 to .842. After inspection of the modification indices, I respecified the model to include the covariance between Item 3 and Item 4, as well as the covariance between Item 7 and Item 8. Wordings for each item are presented in Appendix A. After respecifying the model, model fit improved towards acceptable (RMSEA = .116 [.102, .130], CFI = .924, TLI = .891, SRMR = .051). The Chi-Square test of model fit remained significant (χ^2 (25) = 178.712, p = .000).

For the measurement model specified for timepoint two (T2) *diversity climate*, model fit was also mediocre (RMSEA = .163 [.150, .176], CFI = .844, TLI = .793, SRMR = .069). The Chi-Square test of model fit was significant (χ^2 (27) = 357.479, p = .000). Standardized factor loadings ranged from .59 to .87. I also respecified the measurement model for T2 *diversity climate* to include the covariances between Item 3 and Item 4, as well as Item 7 and Item 8. After respecifying the model, model fit improved such that model fit approached acceptable (RMSEA

= .116, [.102, .130], CFI = .928; TLI = .896; SRMR = .056). The Chi-Square test of model fit remained significant (χ^2 (25) = 178.775, p = .000).

Psychological safety. The measurement model specifying a single factor for *psychological safety* (T1) resulted in mediocre fit (RMSEA = .089 [.065, .115], CFI = .903; TLI = .839; SRMR = .050). The Chi-Square test of model fit was significant (χ^2 (9) = 42.312, p = .000). The standardized factor loadings ranged from .370 to .699. After evaluating modification indices, I chose to respecify the measurement model to include the covariance between Item 1 and Item 3. These items are presented in Appendix A. After respecifying the model, results indicated excellent model fit (RMSEA = .050 [.019, .080]; CFI = .973; TLI = .950; SRMR = .032). The Chi-Square test of model fit test remained significant but improved (χ^2 (8) = 17.264, p = .027).

For T2, the single factor measurement model resulted in mediocre fit (RMSEA = .124 [.099, .151], CFI = .849; TLI = .748; SRMR = .060). The Chi-Square goodness-of-fit test was significant (χ^2 (15) = 73.697, p = .000). Standardized factor loadings ranged from .415 to .700. After respecifying the model to include the covariance between Item 1 and Item 3 the model resulted in acceptable fit (RMSEA = .072 [.044, .102], CFI = .955 TLI = .915; SRMR = .037). The Chi-Square test of model fit remained significant (χ^2 (8) = 27.371, p = .001).

Organizational identity. The single factor measurement model specified for *organizational identity* (T1) resulted in acceptable fit (RMSEA = .077 [.056, .099], CFI = .975, TLI = .958, SRMR = .029). The Chi-Square test of model fit was significant (χ^2 (9) = 33.930, p = .000). The standardized factor loadings ranged from .690 to .888. For T2, the measurement model also resulted in acceptable fit (RMSEA = .068 [.046, .092], CFI = .984; TLI = .973; SRMR = .021). The Chi-Square goodness-of-fit test was significant (χ^2 (9) = 28.559, p = .001).

Safety voice. The two-factor (i.e., *promotive* and *prohibitive*) measurement model for *safety voice* (T1) resulted in acceptable fit (RMSEA = .072 [.049, .097]; CFI = .976; TLI = .955; SRMR = .041). The Chi-Square test of model fit was significant (χ^2 (8) = 27.379, p = .001). Standardized factor loadings for the first factor ranged from .857 to .906, and .674 to .769 for the second factor. For *safety voice* (T2), the measurement model resulted in excellent fit (RMSEA = .064 [.039, .091]; CFI = .982; TLI = .971; SRMR = .026). The Chi-Square test of model fit was significant (χ^2 (8) = 23.353, p = .003).

Safety climate. The single factor measurement model for *safety climate* (T1) resulted in acceptable fit (RMSEA = .081 [.060, .103]; CFI = .962; TLI = .937; SRMR = .034). The Chi-Square test of model fit was significant (χ^2 (9) = 36.290, p = .000). Standardized factor loadings ranged from .664 to .762. For *safety climate* (T2), the measurement model resulted in excellent fit (RMSEA = .057 [.035, .079], CFI = .984; TLI = .973; SRMR = .026). The Chi-Square test of model fit was also significant (χ^2 (9) = 22.426, p = .008).

Between-Groups Invariance Testing

For each variable, I conducted between-groups invariance testing using data collected at T1. I report the from these tests results by variable, including individual group (i.e., White, Black, Hispanic) and multi-group model fit indices and model comparison estimates in Tables 9 through 11.

Diversity climate. The configural model for *diversity climate* approached acceptable fit (RMSEA = .112 [.097, .127]; CFI = .933; TLI = .903; SRMR = .047). The Chi-Square test of model fit was significant (χ^2 (75) = 217.521, p = .000). Invariance testing resulted in metric invariance, indicating the factor structure and loadings were invariant by racial-ethnicity, but the intercepts vary.

Psychological safety. The configural model for *psychological safety* resulted in excellent fit (RMSEA = .039 [.000, .078], CFI = .984; TLI = .970; SRMR = .038). The Chi-Square goodness-of-fit test was non-significant (χ^2 (24) = 29.829, p = .190). Invariance testing resulted in metric invariance, indicating the factor structure and intercepts were equivalent across racial-ethnic groups but the loadings and intercepts vary.

Organizational identity. The configural model for *organizational identity* resulted in adequate fit. (RMSEA = .099, [.075, .123]; CFI = .961; TLI = .936; SRMR = .032). The Chi-Square test of model fit was significant (χ^2 (27) = 67.853, p = .000). Invariance testing resulted in scalar invariance, such that the factor structure, loadings, and intercepts do not vary significantly across racial-ethnic group.

Safety voice. The configural model for safety voice resulted in adequate model fit (RMSEA = .092, [.064, .121], CFI = .966, TLI = .936, SRMR = .043). The Chi-Square test of model fit was significant (χ^2 (24) = 55.710, p = .000). Invariance tests resulted in scalar invariance, such that factor structure, loadings, and intercepts do not vary significantly across racial ethnic groups.

Safety climate. The configural model for safety climate resulted in adequate model fit (RMSEA = .082 [.057, .108]; CFI = .963; TLI = .938; SRMR = .036). The Chi-Square test of model fit was significant (χ^2 (27) = 55.465, p = .001). Invariance tests resulted in scalar invariance, indicating the factor structure, loadings, and intercepts do not vary significantly across racial-ethnic groups.

Longitudinal Invariance Testing

I report results from the longitudinal invariance tests by variable, including model fit indices and model comparison estimates in Tables 12 and 13.

Diversity climate. The configural model specified to test for configural invariance for T1 and T2 *diversity climate* resulted in mediocre fit (RMSEA = .106, [.099, .113], CFI = .871, TLI = .842, SRMR = .062). The Chi-Square test of model fit was significant (χ^2 (125) = 784.363, p = .000). These results indicate that the factor structure specified for *diversity climate* was not equivalent between T1 and T2. Given only T1 *diversity climate* is used within the model for hypothesis testing, this variable was retained for the final model.

Psychological safety. The model specified to test for configural invariance for T1 and T2 psychological safety resulted in acceptable fit (RMSEA = .062, [.049, .074], CFI = .946, TLI = .924, SRMR = 051). The Chi-Square test of model fit was significant (χ^2 (47) = 131.091, p = .000). Invariance tests resulted in scalar invariance, indicating the factor structure, loadings, and intercepts do not vary across timepoints.

Organizational identity. The model specified to test for configural invariance for T1 and T2 organizational identity resulted in excellent fit (RMSEA = .045, [.031, .059], CFI = .986, TLI = .981, SRMR = .024). The Chi-Square test of model fit was significant (χ^2 (47) = 92.014, p = .000). Invariance tests resulted in scalar invariance such that the factor structure, loadings, and intercepts do not vary across timepoints.

Safety Voice. The model specified to test for configural invariance for T1 and T2 safety voice resulted in excellent fit (RMSEA = .038, [.022, .053], CFI = .987, TLI = .979, SRMR = .044). The Chi-Square test of model fit was significant (χ^2 (42) = 70.907, p = .003). Additional invariance testing indicated the model only fit criteria for scalar invariance, such that the factor structure was invariant across factor structure, loadings, and intercepts.

Safety Climate. The model specified to test for configural invariance for T1 and T2 safety climate resulted in excellent fit (RMSEA = .053, [.040, .066], CFI = .969, TLI = 9.56,

SRMR = .031). The Chi-Square test of model fit was significant (χ^2 (47) = 108.459, p = .000). Invariance testing stopped at configural invariance, indicating only the factor structure was invariant across time points.

Hypothesis Testing

For hypothesis testing, I began by testing the hypothesized model without the inclusion of control variables. Next, I rerun my analysis including control variables, but only when variables for consideration were significantly correlated with both the predictor and outcome variables. This approach is in line with recommendations by Carlson and Wu (2012) regarding the conservative use of control variables. Given *safety climate* was the only measured variable that was moderately and significantly correlated with both predictor and outcome variables, I included *safety climate* as a control. Specifically, I included the direct path of safety climate (T1) as predicting the outcome variable of safety voice (T2)

In the following sections, I report results of the controlled model. When running the results without the inclusion of controls, there were minimum differences in the effect size estimates and no difference in the significance of effects. For transparency, direct effects from both the controlled and uncontrolled model are provided in Tables 14 and 15. All results are reported as standardized estimates.

Model Fit. The first execution of the hypothesized model resulted in 5600 iterations and a PSR of 1.105, just above the acceptable criterion for precise convergence. When executing the analysis a second time and increasing the number of iterations to 5800, the PSR decreased to 1.074. This estimate is within the acceptable range for determining precise convergence when using Bayesian estimation (Zittman & Hecht, 2019).

Hypothesis 1. H1 stated diversity climate would positively and significantly predict safety voice. To test H1, I specified diversity climate (T1) as predicting safety voice (T2). The direct effect of diversity climate (T1) on safety voice (T2) was non-significant (b = -.007, [-.100, .080], indicating a lack of support for H1. I report all other direct effects in Table 11.

Hypothesis 2. H2 stated the variables of psychological safety and organizational identity would mediate the relationship between diversity climate and safety voice. To test H2, I began by specifying the a-paths, or the direct paths between diversity climate (T1) and psychological safety (T2) and the direct path between diversity climate (T1) and organizational identity (T2). Next, I specified the b-paths, or the direct paths between psychological safety (T1) and safety voice (T2) and organizational identity (T1) and safety voice (T2). Mediation, or the indirect effects, were tested by taking the product term of separately multiplying the a- and b-paths by variable (i.e., psychological safety and organizational identity).

First, the indirect effect of diversity climate on safety voice via psychological safety was significant (b = .028, [.006, .055]), such that individuals reporting greater levels of diversity climate also reported higher levels of psychological safety, and subsequently higher levels of safety voice. Next, the indirect effect of diversity climate on safety voice via organizational identity was non-significant (b = .009, [-.002, .024]). These results suggest partial support for H2.

Hypothesis 3. H3 stated the indirect effects of diversity climate on safety voice via psychological safety and organizational identity would be the strongest for participants holding a racial-ethnic minority identity, with non-Hispanic Black participants demonstrating the strongest effects, followed by Hispanic, and non-Hispanic White participants, respectively. To test H3, I

specified the interaction term of diversity climate by racial-ethnic group. Racial-ethnicity was dummy coded such that each minority racial-ethnic group was compared to the non-Hispanic White participant group. Next, I specified the interaction terms as predicting both psychological safety and organizational identity. These estimates were then used to test moderated mediation.

Results demonstrated that the indirect effect of diversity climate on safety voice via psychological safety was significant and the strongest for Black as compared to White participants (b = .029, [.009, .057]), but non-significant for Hispanic as compared to White participants (b = .025, [-.002, .068]). Subsequently, the indirect effect of diversity climate on safety voice via organizational identity was non-significant for both Black as compared to White participants (b = .020, [-.015, .065]), and Hispanic as compared to White participants (b = .022, [-.015, .074]). In combination, these results suggest partial support for H3.

Hypothesis 4. H4 stated the indirect effect of diversity climate on safety voice via organizational identification and psychological safety would also lend to differences in reported occupational incidents by racial ethnicity. To test H4, I added the path of safety voice (T1) as predicting the number of reported incidents (T2) to the overall hypothesized model and used this estimate to test moderated serial-mediation. The direct effect of safety voice (T1) on reported incidents (T2) was non-significant (b = .05, [-.005, .105]). The indirect effect of diversity climate via psychological safety and organizational identity on safety voice, and subsequently on reported incidents, was also non-significant across all groups (White = .001, [-.004, .007]; Black = .002, [-.003, .009]; Hispanic = .002, [-.004, .009]). H4 is not supported. I discuss the implications of all results in the following section.

DISCUSSION

In the U.S., minority racial and ethnic groups acquire an undue burden of work-related injury and disability. Yet, beyond documentation that these disparities exist, there is a dearth of empirical research lending to why they occur. In response, I proposed the current study to approach a preliminary understanding of mechanisms which might lend to these occupational health disparities. Specifically, I sought to investigate whether organizational diversity climate perceptions may relate to racial/ethnic differences in worker safety voice via psychological safety and organizational identification, and subsequently differences in occupational incidents. Study results partially supported my hypotheses.

Specifically, diversity climate was shown to have an indirect effect on safety voice via psychological safety, and this effect was strongest for some minority racial-ethnic participants (i.e., Black participants). This effect was not found when considering organizational identification as a mediator, nor did the indirect effect extend to predict reported incidents. In the following sections, I discuss how these findings fit within the present research and respective implications for theory, followed by a discussion of the implications for organizational practice, and limitations to guide future research.

Theoretical Implications

The Indirect Association Between Diversity Climate and Safety Voice

Diversity climate did not directly predict reported safety voice in the current study, but the significant indirect effect from diversity climate to safety voice via psychological safety speaks to the central tenets of Social Identity Theory (Tajfel & Turner, 1985; SIT). Under SIT, individuals perceiving pro-diversity climates should be more likely to adhere to social

identification and engage in behaviors which support and embody the organization. Psychological safety is often modeled as an outcome of pro-diversity climates that is indicative of the social identification process (Ashforth & Mael, 2002). In the current study, participants reporting higher levels of diversity climate perceptions also reported greater levels of psychological safety, and subsequently greater levels of safety voice behaviors. These results further speak to the presumed importance of the social identification process for influencing citizenship behaviors such as safety voice.

The current results also reinforce modern propositions that safety in the workplace is socially constructed (Turner & Gray, 2009). For example, the direct effects of psychological safety and organizational identity on safety voice were both positive and significant (See Table 14), indicating affective reactions associated with socially rooted phenomena are important for influencing important behaviors. However, to advance the theoretical connection between diversity and safety, future research is needed to understand why diversity climate was not predictive of organizational identity in the current study.

One plausible explanation could relate to study sample size and an inefficient power to detect small effects. For the current study, an a priori power analysis suggested a minimum sample requirement of N = 550 (B = .80). After data collection and removing outliers from my sample I ended with a final sample of N = 493. In the current study, there was a relatively small, non-significant effect of diversity climate on organizational identity (b = .07, [-.017, .139], as well as credibility intervals which approached zero. This effect may be determined to be significant if retested within an appropriately powered study, and future research should pursue this hypothesis via additional studies with larger samples.

Racial-Ethnic Differences in Safety Voice Behaviors

Hypothesis 3 (H3) predicted the indirect effect of diversity climate on safety voice would be stronger for racial/ethnic minorities, with non-Black Hispanic workers demonstrating the strongest effects. This hypothesis was partially supported, such that the indirect effect of diversity climate on safety voice, via psychological safety, was significant and strongest for Black participants. In contrast to the hypothesized effect, this indirect effect was also weakest, and non-significant for Hispanic participants. Given SIT assumptions, one would expect both Black and Hispanic participants to demonstrate the strongest effects given their minoritized racial-ethnic identities and the supposed increased influence of diversity climate on affective reactions for minoritized group members (Tajfel & Turner, 1978).

One potential explanation for these unforeseen differences in effects is the difference in mean age between the Hispanic (32.9) versus non-Hispanic White (40.4) and non-Hispanic Black groups (40.7). Hispanic participants also had a lower mean tenure (5.9 years) as compared to non-Hispanic White (7.9 years) and non-Hispanic Black (6.5 years) participants. When evaluating descriptive statistics, these variables also held a significant correlation with T2 *safety voice* (Age: $r = .197$, $p = .009$; Tenure: $r = .230$, $p = .100$). This pattern of results between age, tenure, and safety voice follows empirical research suggesting overall experience as positively correlated with voice behaviors, generally (Whiting et al., 2012). Other research shows younger workers are also less likely to speak up regarding safety concerns, rather due to vulnerability or a lessened understanding of safety policies, procedures, or their rights as workers (Tucker & Turner, 2015). Despite the significant correlations between age, tenure, and safety voice, I did not consider *age* and *tenure* as control variables within models for hypothesis testing due to their non-significant correlation with the predictor variable of *diversity climate* (Age: $r = .05$; Tenure:

r = .05). Nonetheless, future research should consider these variables as moderators or controls in the relationship between social identification and the likelihood to enact safety voice behaviors.

An alternative approach for evaluating this relationship is through a lens of intersectionality. The concept of intersectionality gives attention to the interaction between multiple intersecting identities (e.g., race, ethnicity, gender, age, etc.), their ascribed social positions, and the perceived power associated with those positions (Rosenthal, 2016). Given the response patterns discussed above, future research may also look at the effects of holding multiple minoritized or vulnerable identities on the likelihood to engage in safety voice behaviors. For example, Hispanic workers may experience vulnerability in the workplace due to their minoritized identity, but workers who are both Hispanic and young may experience even greater vulnerability and lesser likelihood to speak up on safety concerns. A better understanding of how individual intersectionality relates to safety voice will likely strengthen the theoretical bridging of diversity and safety literatures.

Racial Ethnic Difference in Incident Occurrence

In contrast to H4, the current study did not demonstrate a serial effect linking safety voice to incident occurrence. However, descriptive statistics revealed there is a significant relationship between identifying as a racial-ethnic minority and having reported experiencing a workplace incident (r = .107, p = .022). One potential explanation for the lack of significant effects in the current study may be due to a lack of systems considerations. For instance, hypothesis development was primarily driven by the use of SIT as a theoretical underpinning, which is helpful for addressing the social phenomenon underlying H1-H3. However, the proposed

relationship between safety voice and incident occurrence as proposed in H4 may likely rely on alternative processes or systems.

For example, speaking up on a safety concern at work does not equate to the concern or hazard being mitigated. Factors such as one's direct supervisor's support for safety directly impact the likelihood of incurring a work-related injury even when an individual demonstrates high levels of safety voice (Tucker & Turner, 2015). From a theoretical perspective, testing the relationship between safety voice and incident occurrence may be better pursued via a sociotechnical systems lens in which takes into consideration multiple layers of a work system including the social, organizational, and external environment (Caryon et al., 2015).

Finally, there are a number of methodological constraints which may have contributed to a lack of effects given H4, which I discuss within the Limitations & Future Directions section of this document. First, however, I will discuss what the current study results mean practically and how organizations may utilize them to better the occupational safety of their diverse workforce.

Implications for Practice

The primary implication for organizations stemming from the current study results is the identification of an additional actionable mechanism for improving the likelihood of worker safety voice, and subsequently occupational safety outcomes. While there is large literature base linking safety climate to outcomes such as safety voice (Sun et al., 2022) and occupational incidents (Griffin & Curcuruto, 2016), the current results reflect diversity climate may also serve as an indirect mechanism for addressing psychosocial processes lending to important safety outcomes. Organizations should be making intentional efforts to enhance their diversity climate in order to bolster safety-related voice behaviors and potentially mitigate differential outcomes

specific to minorities in the workplace.. To do so, organizations should cultivate systems and practices which cue to their workforce they value diversity.

Prior research suggests climate-level outcomes are reflective of individual perceptions based on both formal and informal features of the organization to include policies, procedures, organizational routines, and behaviors that are supported and rewarded within the work environment (Schneider, 1990; Schneider et al., 1996). Additional research has demonstrated the utilization of diversity-specific programs serves as an effective mechanism which cues to employees the organization is committed to and values diversity (Herdman & McMillian-Capehart, 2010). For example, human resource practices (e.g., recruitment and selection) which reflect a fair and consistent diversity objective are one way to strength diversity climate within an organization (Shen et al., 2009). However, considering the current results seem to hinge upon the social processes which lend to organizational identification, simply hiring a diverse workforce is unlikely to lend to improvements. Instead, these processes should focus on the inclusive integration of diverse identities within the social context of the work environment, such as through employee resource groups (ERGs).

ERGs are intra-organizational groups that often join to address a singular focus and attract individuals who identity with the group (Friedman & Craig, 2004). Taking a social-identity approach, Welbourne et al. (2016) suggest that given that ERGs are nested within the organization, membership in an ERG should lend itself two a dual-process of social-identification: one which lends to the social identity of the organization, as well as a social identity with the ERG. For example, McPhee et al. (2017) demonstrate how ERGs can foster a connection between both one's racial/ethnic identity as well as a broader organizational identity. Additionally, as ERGs bring together individuals' of like identities and common interest

(Friedman & Craig, 2004), they should also provide an environment of psychological safety. Thus, ERGs developed to address and attract a diverse range of identities, including those related to racial-ethnic identity, may be an effective way to develop diversity climate within organizations.

Limitations and Future Directions

Study Attrition, Non-Response Bias, and Zero-Inflated Outcomes

Like any study, the current project was vulnerable to several limitations which restrict the generalizability of the current findings. For example, longitudinal research designs are known to place substantial burden on participants given the time delegated to completing assessments across time (Ployhart & Vandenberg, 2010). In the current study, I experienced a 12% attrition rate during the one-month interval following T1. One glaring implication of participant attrition for generalizing results is the likelihood the current study was underpowered to detect the small effect sizes underlying my results. A-priori power analyses suggested recruiting enough participants to reach an overall sample of N = 550 for detecting hypothesized effects; however, after attrition and outlier removal my final sample for analysis was N = 469. Thus, in order to gain clarity on whether the current study results are reflective of insufficient power or an actual lack of effects in the population, future research ought to replicate the current study among a greater sample size.

An additional consideration are the reasons participants did not return to complete their assessments between T1 and T2. For instance, if participants who left the study during this time period were subject to injury or harm during resulting in their ability to complete the study, then the current results may also be subject to concerns related to a non-response or health worker bias, wherein only workers experiencing no, or few, incidents returned to complete the study. To

better understand any potential influence of study attrition on incidents occurred across participants, future research may replicate the current study and assess occupational safety and health history during the first time-point. This process would allow researchers to identify and safety-related characteristic similarities among those participants who did not return for the second timepoint. Along similar lines, Hispanic participants were not only the most difficult to recruit, but also the hardest to retain. Given the small sample size, there is not enough data to make reasonable assumptions regarding the attrition rate for Hispanic participants. However, in the event the data present a healthy worker bias, then group-level differences in incident occurrence by racial-ethnic groups would be severely influenced.

These factors lend to another primary limitation of the current study, wherein very few workers reported experiencing an incident within the one-month interval between T1 and T2 assessments. In fact, the average number of injuries for the entire sample was fewer than one (M = .21, SD = .57). Given the low number of reported incidents, the current results are unlikely to reflect phenomena occurring within the larger population. Additionally, taking near miss count into consideration did not mitigate this design challenge. Across the entire sample, workers also reported a mean near miss count of less than one (M = .40, SD = .87). In order to better evaluate whether racial-ethnic differences in the indirect effect of diversity climate on safety voice also lend to differential incident rates, future researchers ought to replicate the current study over a greater timeframe, such as six months or one-year.

Occupational Characteristics

Another important limitation to the generalizability of the current results is related to the occupational roles held by responding participants. While all participants worked within an overarching industry considered to be high-hazard, not all participants held roles that required

the daily or recurring use of safety procedures, or those wherein their individual job tasks were conducted with high-hazard environments. For example, job titles such as Processing Technician, Primary Care Scheduler, and Operations Support may be exposed to lower levels of risk and hazards. Additionally, participant qualitative reports of their job titles reflected a proportion of workers held management or supervisor roles, wherein the overall proportion of working hours spent in direct exposure to hazards is likely to differ (i.e., time spent in the field versus office). Those looking to replicate the current study would benefit from gathering a sample more concretely exposed to high-hazard job tasks, versus working in a high-hazard industry as a whole.

In the current study I utilized Prolific to recruit participants and was subject to utilizing screening criteria available from the platform (i.e., screening participants based on whether they considered safety procedures within their work). For more effective recruiting efforts, researchers should consider replicating the current study within a singular organization. Partnering with an organization may serve several additional benefits, including *a)* the potential to evaluate safety incidents which have occurred across a greater time frame, including those prior to conducting the study, *b)* better target participants who have the greater potential for occupational incidents for inclusion, and *c)* evaluate the impact of diversity climate on safety outcomes at multiple levels of personnel (i.e., staff, supervisory, executive, etc.).

Further, the analyses in the current study did not take into consideration the proportion of time spent working across participants in relation to the number of incidents reported. Thus, it is possible workers reporting greater incidents also spent more time working during the time between each survey. My analyses also did not incorporate level of risks across the various industries and occupations included. While the screening process was set to identify workers

who were among high-risk industries requiring the use of safety procedures, it is likely level of hazard exposure varied across participants. In order to better generalize study results to the workforce, future research ought to employ studies more concretely measuring these compounding variables.

English as a Second Language

A final consideration is the current study procedures only allowed for English-speaking participants. Relatedly, 13% of the United States population are Spanish-only speakers (Census.gov, 2022). Considering this metric, it is likely a large body of Hispanic participants were excluded from the current study. Notably, Hispanic and English as a second language (ESL) workers have been shown to report greater levels of work-related pain lending to workdays missed (Pemji & Krause, 2010) and work in roles with greater occupational risks (Davila et al., 2011). Given a primary variable in the current study is related to the communication of safety-related information meant to reduce incidents and injury, the exclusion of a portion of racial/ethnic minority workers severely influences the generalizability of the current study. To better understand how racial/ethnic disparities in occupational illness and injury occur, and subsequently can be mitigated, future research must include ESL workers.

CONCLUSION

In the current study, I sought to empirically examine racial-ethnic differences in the relationship between diversity climate and safety voice, and whether these differences were subsequently related to occupational health burdens experienced by racial-ethnic minorities. Results suggest individual diversity climate perceptions serve a modest role for encouraging worker safety voice behaviors, primarily due to increases in psychological safety among workers perceiving pro-diversity climates. Workers who perceived more positive diversity climates, also reported higher rates of psychological safety, and subsequently greater rates of safety voice. These results were also shown to be the strongest among Black-identifying racial-ethnic minorities. However, study results did not indicate support for whether diversity climate perceptions eventually lend to differential rates in occupational incidents by racial-ethnicity, and further research is needed. Nonetheless, the current results do speak to a preliminary mechanism for targeting occupational health disparities via diversity climate, especially given the rich history of evidence supporting the connection between safety voice and reduced occupational incidents. Thus, organizations may refer to the current results as a preliminary cue for paying attention to mechanisms for supporting the diverse characteristics of their workforce, in order to better the occupational safety of their workers regardless of identity.

TABLES AND FIGURES

Table 1

Participant Demographics

	M	***SD***
Age [485]	*38.62*	*11.07*
Gender	**N**	**%**
Man	295	60.0
Woman	183	37.1
Transgender Man	8	1.6
Other	4	0.8
Did Not Specify	2	0.4
Transgender Woman	1	0.2
Racial Ethnicity [492]	**N**	**%**
Non-Hispanic White	194	39.6
Non-Hispanic Black	176	35.7
Hispanic	123	24.7
Tenure in Years [489]	***M***	***SD***
	6.9	5.9
Occupational Industry	**N**	**%**
Health Services	134	27.18%
Manufacturing	66	13.39%
Other [a]	63	12.79%
Retail Trade	60	12.17%
Construction	59	11.97%
Leisure and Hospitality	29	5.88%
Transportation	28	5.68%
Warehousing/Storage	23	4.67%
Food Processing	14	2.84%
Agriculture/Fishing/Hunting	10	2.03%
Utilities	4	0.81%
Forestry and Logging	1	0.20%
Mining	1	0.20%
Waste Management	1	0.20%

Note. All proportions are calculated for N = 493 unless otherwise specified in [brackets].
[a] Occupational industries included in "Other", included Information Technology, Military and Protective Services, Research and Science, Education, and Social Work.

Table 2

Non-Hispanic White Participant Demographics (N = 194)

	M	***SD***
Age	*40.37*	*11.24*
Gender	**N**	**%**
Man	124	64.6
Woman	60	32.8
Transgender Woman	2	1.0
Other	2	1.0
Did Not Specify	1	0.5
Tenure in Years	***M***	***SD***
	7.9	6.6
Occupational Industry	**N**	**%**
Health Services	52	26.7%
Manufacturing	34	17.4%
Retail Trade	27	13.9%
Construction	24	12.3%
Other	22	11.3%
Warehousing/Storage	7	3.6%
Leisure and Hospitality	6	3.1%
Food Processing	5	2.6%
Agriculture/Fishing/Hunting	5	2.6%
Utilities	4	2.0%
Transportation	2	1.0%
Mining	1	0.5%
Forestry and Logging	0	0.0%
Waste Management	0	0.0%

Table 3

Non-Hispanic Black Participant Demographics (N = 176)

	M	***SD***
Age	*40.68*	*11.03*
Gender	**N**	**%**
Man	97	55.1%
Woman	72	52.3%
Transgender Man	6	3.4%
Other	1	0.6%
Tenure in Years	***M***	***SD***
	6.5	5.7
Occupational Industry	**N**	**%**
Health Services	43	24.4%
Construction	27	15.3%
Other	27	15.3%
Manufacturing	19	10.80%
Retail Trade	17	9.7%
Transportation	14	8.0%
Leisure and Hospitality	11	6.3%
Warehousing/Storage	7	4.0%
Food Processing	5	2.8%
Agriculture/Fishing/Hunting	3	1.7%
Utilities	2	1.1%
Forestry and Logging	1	0.6%
Mining	0	0.0%
Waste Management	0	0.0%

Table 4

Hispanic Participant Demographics (N = 122)

	M	***SD***
Age	*32.9*	*8.7*
Gender	**N**	**%**
Man	72	59.0%
Woman	47	38.5%
Transgender Man	2	1.6%
Other	1	0.8%
Tenure in Years	***M***	***SD***
	5.8	4.8
Occupational Industry	**N**	**%**
Health Services	39	31.97%
Retail Trade	16	13.11%
Other	13	10.66%
Manufacturing	13	10.66%
Construction	9	7.38%
Leisure and Hospitality	9	7.38%
Warehousing/Storage	9	7.38%
Transportation	7	5.74%
Food Processing	4	3.28%
Agriculture/Fishing/Hunting	2	1.64%
Waste Management	1	0.82%
Utilities	0	0.00%
Forestry and Logging	0	0.00%
Mining	0	0.00%

Table 5

Descriptive Statistics for T1 and T2 Variables

Variable	*M*	*SD*	Skew	Kurtosis	SE
Diversity Climate T1	3.62	0.92	-0.48	-0.37	0.04
Diversity Climate T2	3.60	0.92	-0.39	-0.50	0.04
Psych Safety T1	3.73	0.73	-0.28	-0.29	0.03
Psych Safety T2	3.73	0.76	-0.43	-0.10	0.04
Org. Identification T1	3.43	1.14	-0.50	-0.75	0.05
Org. Identification T2	3.38	1.17	-0.43	-0.86	0.05
Safety Voice T1	3.98	0.87	-1.04	0.78	0.04
Safety Voice T2	3.97	0.78	-0.79	0.30	0.04
Safety Climate T1	4.10	0.79	-0.92	0.50	0.04
Safety Climate T2	4.12	0.72	-0.54	-0.63	0.03
Incident Number	0.20	0.57	3.80	18.79	0.03
Near Miss Number	0.40	0.87	2.94	10.07	0.04
Age (years)	38.86	11.14	0.53	-0.37	0.53
Tenure (years)	7.03	5.97	1.61	3.16	0.28
Gender	0.71	0.75	2.61	11.96	0.04

Note. T1 = Timepoint 1. T2 = Timepoint 2. Org. Identification = Organizational Identification. Gender coded 1 = Male, 0 = Female.

Table 6

Correlations Among all T1 and T2 Variables

	Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	Div 1	-														
2	Div 2	.75	-													
3	Psy 1	.38	.37	-												
4	Psy 2	.38	.41	.76	-											
5	Org 1	.45	.42	.33	.39	-										
6	Org 2	.42	.46	.31	.42	.88	-									
7	SafV 1	.35	.30	.28	.30	.45	.42	-								
8	SafV 2	.35	.35	.37	.42	.45	.45	.74	-							
9	SafC 1	.45	.38	.44	.45	.48	.40	.45	.39	-						
10	SafC 2	.42	.49	.49	.54	.40	.41	.42	.51	.67	-					
11	Inc	.07	.06	-.11	-.03	.15	.15	.06	.10	-.04	-.04	-				
12	NM	.10	.14	.02	.00	.21	.21	.16	.16	.05	.06	.38	-			
13	Age	.06	.07	.01	.07	.20	.25	.19	.20	.14	.19	.13	.12	-		
14	Tenure	.05	.05	.11	.12	.19	.21	.25	.22	.19	.18	.03	.00	.49	-	
15	Gender	.06	.03	.13	.11	.01	.00	.00	.03	.00	.05	-.02	.02	-.02	.00	-

Note. Div 1 = Diversity Climate Timepoint 1. Div 2 = Diversity Climate Timepoint 2. Psy 1 = Psychological Safety Timepoint 1. Psy 2 = Psychological Safety Timepoint 2. Org 1 = Organizational Identity Timepoint 1. Org 2 = Organizational Identity Timepoint 2. SafV 1 = Safety Voice Timepoint 1. SafV 2 = Safety Voice Timepoint 2. SafC 1 = Safety Climate Timepoint 1. SafC 2 = Safety Climate Timepoint 2. Inc = Incidents Occurred. NM = Near Misses Occurred. Gender (0 = Female, Male = 1).

Table 7

Incident Descriptives by Race/Ethnicity

Race/Ethnicity	N	Total Incidents	Mean Incidents	*SD*	Total Near Miss	Mean Near Miss	*SD*
All Groups	492	93	.20	.57	175	.40	.87
Non-Hispanic White	194	23	.12	.41	43	.24	.69
Non-Hispanic Black	176	47	.28	.73	90	.59	1.06
Hispanic	122	23	.2	.52	42	.24	.78

Table 8

Scale Fit Indices and Standardized Reliability (Alpha) Estimates by Timepoint

Variable	Time	χ^2	CFI	TLI	RMSEA	SRMR	*a*
Diversity Climate	T1	178.712*	.924	.891	.116	.051	.931
	T2	178.775*	.928	.896	.116	.056	.931
Psychological Safety	T1	17.264*	.973	.950	.050	.032	.711
	T2	27.371*	.955	.915	.072	.037	.745
Org. Identification	T1	33.930*	.975	958	.077	.029	.910
	T2	28.559*	.984	.973	.068	.021	.925
Safety Voice	T1	27.379*	.976	.955	.072	.041	.873
	T2	23.353*	.982	.971	.064	.026	.856
Safety Climate	T1	36.290*	.962	.937	.081	.034	.870
	T2	22.426*	.984	.973	.057	.026	.897

Note. Some items among latent variables were allowed to covary given similarity of wording and response patterns. Org. Identification = Organizational Identification. T1 = Timepoint 1. T2 = Timepoint 2. Chi-square estimates marked with an asterisk (*) indicates $p < .05$.

Table 9

Fit Indices and Chi-Square Estimates by Variable and Racial-Ethnicity

Variable	Racial-Ethnicity	χ^2	CFI	TLI	RMSEA	SRMR
Diversity Climate	White	117.872*	.899	.855	.139	.069
	Black	34.480*	.983	.976	.057	.030
	Hispanic	81.407*	.909	.869	.137	.047
Psychological Safety	White	9.262	.994	.990	.029	.030
	Black	11.554	.969	.942	.050	.049
	Hispanic	13.401	.957	.920	.074	.055
Org. Identification	White	43.368*	.933	.888	.140	.042
	Black	9.588	.999	.998	.019	.021
	Hispanic	14.922	.977	.961	.073	.038
Safety Voice	White	34.782*	.948	.903	.131	.053
	Black	14.278	.980	.952	.068	.052
	Hispanic	8.199	.999	.998	.014	.031
Safety Climate	White	15.447	.987	.979	.061	.026
	Black	11.916	.993	.989	.043	.023
	Hispanic	29.523*	.924	.873	.138	.050

Note. Some items among latent variables were allowed to covary given similarity of wording and response patterns. Org. Identification = Organizational Identification. Chi-square estimates marked with an asterisk (*) indicates $p < .05$.

Table 10

Between-Groups Invariance Testing Model Fit Indices by Variable

Variable	Model	χ^2	CFI	TLI	RMSEA	SRMR
Diversity Climate	Configural	217.521*	.933	.903	.113	.047
	Metric	247.153*	.926	.912	.106	.070
	Scalar	320.525*	.899	.898	.114	.083
Psychological Safety	Configural	29.829	.984	.970	.039	.038
	Metric	48.000	.961	.949	.051	.062
	Scalar	72.666*	.921	.919	.065	.075
Organizational Identification	Configural	67.853*	.961	.936	.099	.032
	Metric	79.073*	.960	.952	.085	.043
	Scalar	90.608*	.959	.961	.077	.046
Safety Voice	Configural	55.710*	.966	.936	.092	.043
	Metric	65.817*	.964	.949	.083	.065
	Scalar	75.708*	.962	.957	.076	.066
Safety Climate	Configural	55.465*	.963	.938	.082	.036
	Metric	64.773*	.964	.956	.070	.059
	Scalar	78.884*	.959	.960	.066	.063

Table 11

Between-Groups Invariance Testing Model Comparison

Variable	Models Compared	χ^2	*p*
Diversity Climate	Configural / Metric	305.51	.055
	Metric / Scalar	383.19	.000
Psychological Safety	Configural / Metric	81.746	.216
	Metric / Scalar	106.702	.004
Organizational Identification	Configural / Metric	105.809	.804
	Metric / Scalar	113.714	.552
Safety Voice	Configural / Metric	90.905	.153
	Metric / Scalar	98.701	.449
Safety Climate	Configural / Metric	95.076	.499
	Metric / Scalar	108.085	.230

Table 12

Longitudinal Invariance Testing Model Fit Indices by Variable

Variable	Model	χ^2	CFI	TLI	RMSEA	SRMR
Diversity Climate	Configural	784.363*	.871	.842	.106	.062
	Metric	-	-	-	-	-
	Scalar	-	-	-	-	-
Psychological Safety	Configural	131.091*	.946	.924	.062	.051
	Metric	120.138*	.965	.956	.053	.045
	Scalar	143.455*	.944	.936	.057	.053
Organizational Identification	Configural	92.014*	.986	.981	.045	.024
	Metric	99.593*	.985	.982	.044	.032
	Scalar	105.285*	.985	.983	.042	.033
Safety Voice	Configural	70.907*	.987	.971	.038	.044
	Metric	107.078*	.973	.962	.051	.090
	Scalar	153.925*	.971	.964	.064	.090
Safety Climate	Configural	108.495*	.969	.956	.053	.031
	Metric	120.138*	.965	.956	.053	.045
	Scalar	125.273*	.964	.957	.052	.047

Note. No estimates are provided for Diversity Climate for Metric and Scalar invariance due to not meeting requirements for configural invariance.

Table 13

Longitudinal Invariance Testing Model Comparison

Variable	Models Compared	χ^2	*p*
Diversity Climate	Configural / Metric	-	-
	Metric / Scalar	-	-
Psychological Safety	Configural / Metric	1.378	.927
	Metric / Scalar	8.606	.126
Organizational Identification	Configural / Metric	6.528	.258
	Metric / Scalar	4.771	.445
Safety Voice	Configural / Metric	35.790	.000
	Metric / Scalar	-	-
Safety Climate	Configural / Metric	11.721	.039
	Metric / Scalar	-	-

Note. No comparisons conducted for diversity climate due to stopping invariance testing after configural invariance criteria were unmet.

Table 14

Direct Effects Estimates and Lower and Upper 95% Credibility Intervals from the Hypothesized Model Specified With Control Variables

	b	Lower 95%	Upper 95%
Psychological Safety (T2) on			
Psychological Safety (T1)	.911	.867	.949
Diversity Climate (T1)	.168	.045	.310
Organizational Identification (T2) on			
Organizational Identity (T1)	.926	.901	.948
Diversity Climate (T1)	.066	-.017	.154
Safety Voice (T2) on			
Safety Voice (T1)	.739	.668	.803
Safety Climate (T2)	.092	-.100	.106
Diversity Climate (T1)	-.007	-.100	.080
Psychological Safety (T1)	.178	.082	.275
Org. Identification (T1)	.143	.052	.235
Incident Number (T2) on			
Safety Voice (T1)	.096	-.005	.104

Note. T1 = Timepoint 1. T2 = Timepoint 2. Org. Identification = Organizational Identification. "On" synonymous with "predicted by".

Table 15

Direct Effects Estimates and Lower and Upper 95% Credibility Intervals from the Hypothesized Model Specified Without Control Variables

	b	Lower 95%	Upper 95%
Psychological Safety (T2) on			
Psychological Safety (T1)	.909	.869	.948
Diversity Climate (T1)	.174	.047	.299
Organizational Identification (T2) on			
Organizational Identity (T1)	.924	.897	.945
Diversity Climate (T1)	.069	-.014	.157
Safety Voice (T2) on			
Safety Voice (T1)	.741	.676	.797
Diversity Climate (T1)	-.004	-.097	.082
Psychological Safety (T1)	.177	.092	.265
Org. Identification (T1)	.143	.056	.235
Incident Number (T2) on			
Safety Voice (T1)	.050	-.004	.107

Note. T1 = Timepoint 1. T2 = Timepoint 2. Org. Identification = Organizational Identification. "On" synonymous with "predicted by".

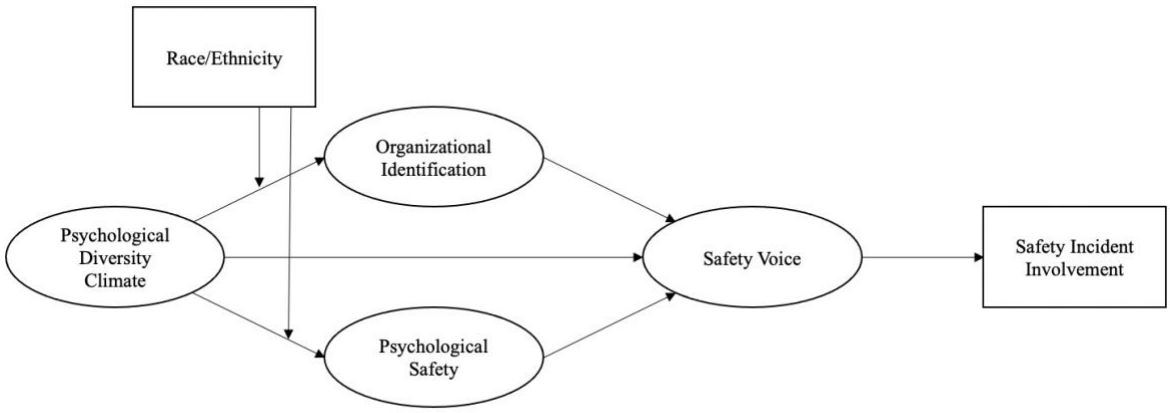

Figure 1

Proposed Direct and Indirect Effects (H1-H4)

REFERENCES

Ashforth, B. E., & Mael, F. (1989). Social identity theory and the organization. *Academy of Management Review*, *14*(1), 20-39.

Barrett, P. (2007). Structural equation modelling: Adjudging model fit. *Personality and Individual Differences*, *42*(5), 815-824.

Bienefeld, N., Grote, G. (2014). Speaking up in ad hoc multiteam systems: Individual-level effects of psychological safety, status, and leadership within and across teams. *European Journal of Work and Organizational Psychology*, *23*(6), 930-945.

Bollen, K. A., Harden, J. J., Ray, S., & Zavisca, J. (2014). BIC and alternative Bayesian information criteria in the selection of structural equation models. *Structural equation modeling: a multidisciplinary journal*, *21*(1), 1-19.

Brown, S., Brooks, R. D., & Dong, X. S. (2021). Injury inequalities among US construction workers. *Journal of Occupational and Environmental Hygiene*, *18*(4-5), 159-168.

Carayon, P., Hancock, P., Leveson, N., Noy, I., Sznelwar, L., & Van Hootegem, G. (2015). Advancing a sociotechnical systems approach to workplace safety–developing the conceptual framework. *Ergonomics*, *58*(4), 548-564.

Census.gov (2022). Nearly 68 Million People Spoke a Language Other Than English at Home in 2019. Accessed on December 22, 2023. https://www.census.gov/library/stories/2022/12/languages-we-speak-in-united-states.html.

Chari, R., Chang, C. C., Sauter, S. L., Sayers, E. L. P., Huang, W., & Fisher, G. G. (2021). NIOSH worker well-being questionnaire (WellBQ).

Cheong, J., MacKinnon, D. P., & Khoo, S. T. (2003). Investigation of mediational processes using parallel process latent growth curve modeling. *Structural Equation Modeling*, *10*(2), 238-262.

Clark, R., Anderson, N. B., Clark, V. R., & Williams, D. R. (1999). Racism as a stressor for African Americans: A biopsychosocial model. *American Psychologist*, *54*(10), 805-816.

Cole, D. A., & Maxwell, S. E. (2003). Testing mediational models with longitudinal data: questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology*, *112*(4), 558.

Collins, L. M., Graham, J. J., & Flaherty, B. P. (1998). An alternative framework for defining mediation. *Multivariate Behavioral Research*, *33*(2), 295-312.

De Cremer, D. (2002). Respect and cooperation in social dilemmas: The importance of feeling included. *Personality and Social Psychological Bulletin*, *28*(10), 1335-1341.

Demerouti, E., Bakker, A. B., Nachreiner, F., & Schaufeli, W. B. (2001). The job demands-resources model of burnout. *Journal of Applied Psychology*, *86*(3), 499-512.

Ely, R. J., & Thomas, D. A. (2001). Cultural diversity at work: The effects of diversity perspectives on work group processes and outcomes. *Administrative Science Quarterly*, *46*(2), 229-273.

Gelman, A., & Rubin, D. B. (1992). Inference rom iterative simulation using multiple sequences. *Statistical Sciences,* 7, 457-472.

Gonzalez, J. A., & DeNisi, A. S. (2009). Cross-level effects of demography and diversity climate on organizational attachment and firm effectiveness. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, *30*(1), 21-40.

Griffin, M. A., & Curcuruto, M. (2016). Safety climate in organizations. *Annual Review of Organizational Psychology and Organizational behavior*, *3*, 191-212.

Hoffmann, D. A., & Morgeson, F. P. (1999). Safety-related behavior as a social exchange: The role of perceived organizational support and leader-member exchange. *Journal of Applied Psychology*, *84*, 286-296.

Hofhuis, J., van Der Zee, K. I., & Otten, S. (2012). Social identity patterns in culturally diverse organizations: the role of diversity climate 1. *Journal of Applied Social Psychology*, *42*(4), 964-989.

Hu, X., & Casey, T. (2021). How and when organization identification promotes safety voice among healthcare professionals. *Journal of Advanced Nursing*, *77*(9), 3733-3744.

Hu, D., Zhang, B., & Wang, M. (2015). A study on the relationship among transformational leadership, organizational identification and voice behavior. *Journal of Service Science and Management*, *8*(01), 142-148.

Islam, T., Ahmed, I., & Ali, G. (2018). Effects of ethical leadership on bullying and voice behavior among nurses: mediating role of organizational identification, poor working condition and workload. *Leadership in Health Services*, *32*(1), 2-17.

Kline, R. B. (2018). Assessing statistical aspects of test fairness with structural equation modelling. In *Fairness Issues in Educational Assessment* (pp. 116-134). Routledge.

Kline, R. B. (2015). The mediation myth. *Basic and Applied Social Psychology*, *37*(4), 202-213.

Landsbergis, P. A., Gryzwacz, J. G., & LaMontagne, A. D. (2014). Work organization, job insecurity, and occupational health disparities. *American Journal of Industrial Medicine*, *57*(5), 495-515.

Liang, J., Farh, C. I., & Farh, J. L. (2012). Psychological antecedents of promotive and prohibitive voice: A two-wave examination. *Academy of Management Journal*, *55*(1), 71-92.

Little, T. D., Preacher, K. J., Selig, J. P., & Card, N. A. (2007). New developments in latent variable panel analyses of longitudinal data. *International Journal of Behavioral Development*, *31*(4), 357-365.

Liu, W., Zhu, R., & Yang, Y. (2010). I warn you because I like you: Voice behavior, employee identifications, and transformational leadership. *The Leadership Quarterly*, *21*(1), 189-202.

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology*, *58*(1), 593-614.

Mael, F., & Ashforth, B. E. (1992). Alumni and their alma mater: A partial test of the reformulated model of organizational identification. *Journal of Organizational Behavior*, *13*(2), 103-123.

McKay, P. F., Avery, D. R., Tonidandel, S., Morris, M. A., Hernandez, M., & Hebl, M. R., (2007). Racial differences in employee retention: Are diversity climate perceptions the key? *Personnel Psychology*, *60*(1), 35-62.

Merritt, S. M., Ryan, A. M., Mack, M. J., Leeds, J. P., & Schmitt, N. (2010). Perceived ingroup and outgroup preference: A longitudinal causal investigation. *Personnel Psychology*, *63*(4), 845-879.

Mor Barak, M. E., Cherin, D. A., & Berkman, S. (1998). Organizational and personal dimensions in diversity climate: Ethnic and gender differences in employee perceptions. *The Journal of Applied Behavioral Science*, *34*(1), 82-104.

Murray, L. R. (2003). Sick and tired of being sick and tired: scientific evidence, methods, and research implications for racial and ethnic disparities in occupational health. *American Journal of Public Health*, *93*(2), 221-226.

Muthén, B., & Asparouhov, T. (2012). Bayesian structural equation modeling: a more flexible representation of substantive theory. *Psychological methods*, *17*(3), 313.

National Institute for Occupational and Health (2019, December 18). *Occupational health equity*. Centers for Disease Control and Prevention https://www.cdc.gov/niosh/programs/ohe/default.html.

National Safety Council (2022). Deaths by demographics. Retrieved from: https://injuryfacts.nsc.org/all-injuries/deaths-by-demographics/race-and-ethnicity/

Premji, S., & Krause, N. (2010). Disparities by ethnicity, language, and immigrant status in occupational health experiences among Las Vegas hotel room cleaners. *American Journal of Industrial Medicine*, *53*(10), 960-975.

Noort, M. C., Reader, T. W., Gillespie, A. (2019). Speaking up to prevent harm: A systematic review of the safety voice literature. *Safety Science*, *117*, 375-387.

Nye, C. D. (2023). Reviewer resources: Confirmatory factor analysis. *Organizational Research Methods*, *26*(4), 608-628.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, 111-163.

Rosenthal, L. (2016). Incorporating intersectionality into psychology: An opportunity to promote social justice and equity. *American Psychologist*, *71*(6), 474.

Ruan, R., & Chen, W. (2021). Relationship between organizational identification and employee voice: A meta-analysis. *Social Behavior and Personality: An Internal Journal, 49*(10), 1-12.

Seabury, S. A., Terp, S., & Boden, L. I. (2017). Racial and ethnic differences in the frequency of workplace injuries and prevalence of work-related disability. *Health Affairs*, *36*(2), 266-273.

Singh, B., Winkel, D. E., & Selvariajan, T. T. (2013). Managing diversity at work: Does psychological safety hold the key to racial differences in employee performance? *Journal of Occupational and Organizational Psychology*, *86*(2), 242-263.

Smith, C. K., Wuellner, S., & Marcum, J. (2023). Racial and ethnic disparities in workers' compensation claims rates. *PloS one*, *18*(1), e0280307.

Stanton, K., Carpenter, R. W., Nance, M., Sturgeon, T., & Villalongo Andino, M. (2022). A multisample demonstration of using the prolific platform for repeated assessment and psychometric substance use research. *Experimental and Clinical Psychopharmacology, 30*(4), 432–443.

Sun, Y., Yang, H., Wu, X., Jiang, Y., & Qian, C. (2022). How safety climate impacts safety voice—Investigating the mediating role of psychological safety from a social cognitive perspective. *International Journal of Environmental Research and Public Health*, *19*(19), 11867.

Tangirala, S., & Ramanujam, R. (2008). Exploring nonlinearity in employee voice: The effects of personal control and organizational identification. *Academy of Management Journal*, *51*(6), 1189-1203.

Tucker, S., Chmiel, N., Turner, N., Hershcovis, M. S., & Stride, C. B. (2008). Perceived organizational support for safety and employee safety voice: the mediating role of coworker support for safety. *Journal of Occupational Health Psychology*, *13*(4), 319-330.

Tucker, S., & Turner, N. (2015). Sometimes it hurts when supervisors don't listen: The antecedents and consequences of safety voice among young workers. *Journal of Occupational Health Psychology*, *20*(1), 72-81.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3*(1), 4-70.

Wang, Y. A., & Rhemtulla, M. (2021). Power analysis for parameter estimation in structural equation modeling: A discussion and tutorial. *Advances in Methods and Practices in Psychological Science*, *4*(1), 1-17.

Wang, Y., Zheng, Y., & Zhu, Y. (2018). How transformational leadership influences employee voice behavior: The roles of psychological capital and organizational identification. *Social Behavior and Personality: An International Journal*, *46*(2), 313-321.

Whiting, S. W., Maynes, T. D., Podsakoff, N. P., & Podsakoff, P. M. (2012). Effects of message, source, and context on evaluations of employee voice behavior. *Journal of Applied Psychology, 97*(1), 159–82. https://doi.org/10.1037/a0024871

Zitzmann, S., & Hecht, M. (2019). Going beyond convergence in Bayesian estimation: Why precision matters too and how to assess it. *Structural Equation Modeling: A Multidisciplinary Journal*, *26*(4), 646-661.

APPENDIX

Appendix A

Scales for Primary Study Variables

Perceptions of Diversity Climate (McKay et al., 2007)
Instructions: To what extent does your company engage in each of the following…
1. Recruiting from diverse sources
2. Offer equal access to training
3. Open communication on diversity
4. Publicize diversity principles
5. Offer training to manage diverse population
6. Respect perspectives of people like me
7. Maintains diversity-friendly work environment
8. Workgroup has climate that values diverse perspective
9. Top leaders visibly committed to diversity
Note. Five-point Likert-type scale (1 = *well below expectations* to 5 = *well above expectations*).

Psychological Safety (Adapted from Edmondson, 1999)
1. If I make a mistake at work, it is often held against me. [R]
2. I am able to bring up problems and tough issues at work.
3. People at my work sometimes reject others for being different. [R]
4. I feel safe to take risks at work.
Note. 5-point Likert response format (1= strongly disagree; 5 = strongly agree). [R] indicates reverse scored items.

Organizational Identification
1. When someone criticizes my organization, it feels like a personal insult.
2. I am very interested in what others think about my organization.
3. When I talk about my organization, I usually say “we” rather than “they”.
4. This organization’s successes are my successes.
5. When someone praises my organization, it feels like a personal compliment.
6. If a story in the media criticized my organization, I would feel embarrassed.
Note. 5-point Likert response format (1= strongly disagree; 5 = strongly agree).

Safety Voice (Adapted from Tucker et al., 2008)

1. I make suggestions about how safety can be improved at work.
2. When my colleague is doing something unsafe, I tell them to stop.
3. I discuss new ways to improve safety with my colleagues.
4. I discuss new ways to improve safety with managers and/or supervisors.
5. I inform my boss when I notice a potential safety hazard at work.
6. I report to my boss when my colleagues break safety rules.

Note. 5-point Likert response format (1= strongly disagree; 5 = strongly agree).

Occupational Injury & Near Miss Occurence

1. In the month since your last survey, did you experience any work-related injuries or illnesses?
2. If yes, how many work-related injuries or illnesses?
3. How would you rate the severity of the worst work-related injury or illness which you acquired in the time since your last survey?
4. In the month since your last survey, have you been involved in a workplace safety incident that did not result in injury or illness (i.e., near miss, close call)?
5. Approximately how many workplace safety incidents (i.e., near miss, close call) have you been involved in since your last survey?

Note. Yes/No and numerical open response formats were used.