6-6-2022

Clarifying and Measuring Inclusive Leadership

Kelly Mason Hamilton
Portland State University

Clarifying and Measuring Inclusive Leadership

by

Kelly Mason Hamilton

A dissertation submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy
in
Applied Psychology

Dissertation Committee:
Larry R. Martinez, Chair
Todd E. Bodner
Leslie B. Hammer
Nicholas A. Smith

Portland State University
2022

Abstract

Many organizations view diversity as a strategic business priority that provides important benefits such as increased creativity and innovation. Research indicates, however, that the potential benefits of diversity cannot be realized without employees feeling a sense of *inclusion*, which involves feeling like one belongs and can be themselves at work. Although scholars acknowledge the important role managers play in fostering inclusion, there remains limited research on specific behaviors they can enact to foster inclusion perceptions in their work groups. Additionally, there is a lack of agreement in the literature about the scope of "inclusive leadership." Historically, scholars viewed inclusive leadership as behaviors that enhance employee voice and foster psychological safety, but new theory has reconceptualized it as behaviors that foster perceptions of inclusion. In this dissertation, I sought to clarify the construct and develop and validate a measure of inclusive leadership (the ILQ) based on new theory (i.e., Perry et al., 2020; Randel et al., 2018), which operationalizes inclusive leadership as a multi-dimensional construct involving behaviors that facilitate belongingness, value uniqueness, and prevent and address mistreatment. To do so, I used Hinkin's (1998) approach: Phase 1, item generation, was completed by 13 subject matter experts; Phase 2, content validity evidence, was provided by 45 working adults and 12 subject matter experts; Phase 3, exploratory factor analysis, was conducted based on responses from 275 working adults; Phase 4, confirmatory factor analysis, was conducted based on responses from 273 working adults; Phase 5, convergent and discriminant validity evidence was provided by

255 working adults; and Phase 6, predictive and incremental validity evidence was provided by 352 working adults. Overall, the results supported the hypothesized factor structure, convergent validity, and predictive utility of the ILQ over and above existing measures of inclusive leadership. Findings provided empirical support for a theoretically grounded, multi-dimensional inclusive leadership construct based on Randel et al.'s (2018) and Perry et al's (2020) frameworks. Finally, results demonstrate the theoretical and practical utility of such a measure.

Dedication

I dedicate this dissertation to my mother, Marilyn. You inspired me from a young age to pursue an impactful career that would serve others. From my earliest years, you have who been a model of persistence, dedication, and resilience. In your own words, you made me a "doer." You and Dad supported and encouraged me in all of my endeavors, no matter how far away from home they took me, and for that I will always be grateful.

Acknowledgements

Completing this dissertation has been a tremendous accomplishment and fulfills a lifetime dream and goal. Starting a Ph.D. program mid-career is not for the faint of heart, but I was looking for a rigorous, transformative experience that would launch me into my next chapter. My time at PSU has given me that and so much more, and I couldn't be more grateful.

I could not have completed the journey without the help and support of so many mentors, colleagues, friends, and family. I would like to first express tremendous gratitude to my advisor, Larry Martinez, who has guided my training over the last five years. Larry, I'm truly amazed by how much growth I've achieved in five years' time, and it would not have been possible without you. From Day 1 you have provided me with all the support, mentorship and opportunities I needed to develop as a scholar and practitioner. As you have said, you'll always be my advisor and I'll always be your student. I am proud to call you my mentor, collaborator and friend.

To my dissertation committee: Larry Martinez, Todd Bodner, Leslie Hammer, and Nicholas Smith, thank you for all of your wise guidance, suggestions, and enthusiasm for this project. Each of your expertise has been invaluable in the development of this manuscript and I am grateful for the opportunity to collaborate with each of you. I also want to acknowledge all of the Department of Psychology professors at Portland State who were a part of my training. I will look back fondly on the rigorous coursework, collaborative faculty, teaching opportunities and opportunities to engage in the larger community through the department's applied research.

To my fellow DASH Lab members, cohort mates, and other grad school peers, I'm grateful for our community. Whether we were collaborating on Quant homework (or commiserating about it on the fifth floor), offering up resources on Slack, or celebrating each other's milestones at Rogue, we knew how to be there for each other when it mattered. I am also deeply grateful for my friends and family for their constant support and encouragement. Whether it was sending care packages in the early years of grad school, checking in on me during the comprehensive exams stage, or helping recruit participants for this research project, your support helped carry me through.

To my spouse, Sarah, you are both my inspiration and my rock. Thank you for supporting me as I pursued my dreams and for being my partner on this adventure and always. It's been a journey for both of us.

Writing this dissertation has been an extremely rewarding learning experience through which I learned so much. I hope this research and its application can make a positive difference in the world.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Many organizations view diversity as a strategic business priority that provides important benefits such as increased creativity, innovation, and competitive advantage (Chrobot-Mason & Aramovich, 2013; Groysberg & Connolly, 2013). Recognizing the potential advantages of having a diverse workforce, organizations have increasingly focused efforts on recruiting higher numbers of racial minorities and women to join their ranks. From a diversity management perspective, these numbers are encouraging because they reflect progress in the business community and society at large in achieving equity for workers from diverse backgrounds. Adding more demographic diversity to an organization, however, does not always lead to the retention, promotion, or full participation of its employees from minority or marginalized groups (Randel et al., 2018; Roberson, 2006). Empirical research strongly supports the assertion that efforts to diversify an organization's demographic composition will not be successful if a positive inclusion climate is not already in place to support the retention of diverse individuals (Hebl & Avery, 2012; Nishi, 2013). Indeed, diversity can actually hinder team performance and lead to an increase in harassment and discrimination (Schneider et al., 2000), intergroup conflict (Jehn et al., 2008), and turnover (McKay et al., 2007) depending on how it influences interpersonal dynamics (Hebl & Avery, 2012; Jackson & Joshi, 2011). For example, when individuals who hold similar levels of power and influence in the organization share multiple attributes such as gender and race (e.g., executive roles filled by White men and administrative roles filled by women of color),

this can create “faultlines” that can negatively impact team creativity and performance due to reduced information sharing and socializing (Jackson & Joshi, 2011).

Empirical evidence suggests that employers may increase the positive consequences of work team diversity by fostering positive inclusion climate perceptions (Chrobot-Mason & Aramovich, 2013; Jackson & Joshi, 2011; Jehn & Bezrukova, 2010), which reflect the degree to which an employee perceives that they are an esteemed member of the work group through experiencing treatment that satisfies their needs for belongingness and uniqueness (Chung et al., 2020; Shore et al., 2011). When employees feel like they belong and can “bring their full selves to work,” inclusion is high. Inclusion climate perceptions have been linked to increased employee well-being, job satisfaction, organizational commitment, organizational citizenship behaviors, job performance, job opportunities and career development, work engagement, creativity, new job involvement, and personal responsibility (Acquavita et al., 2009; Brimhall et al., 2014; Carmeli et al., 2010; Cho & Mor Barak, 2008; Choi et al., 2015; Chung et al., 2020; Findler et al., 2007; Hirak et al., 2012; Mor Barak & Cherin, 1998; Mor Barak et al., 2001; Mor Barak & Levin, 2002) and lower levels of discrimination, harassment, and turnover (Chrobot-Mason & Aramovich, 2013; Mor Barak et al., 2006; Nishii et al., 2010).

Organizational leaders, particularly those who manage work groups, are critical to creating inclusive climates because they make many decisions that impact employees directly (e.g., the allocation of resources) and shape factors of the work environment where employees interact (e.g., creating and/or enforcing policy; Nembhard &

Edmondson, 2006; Nishii, 2013; Perry et al., 2020; Randel et al., 2018; Shore et al., 2011; 2018; Wasserman et al., 2008). Additionally, managers can actively work to prevent exclusion of people from marginalized groups, making them important role models for inclusive and exclusionary behavior (Hollander, 2009; Perry et al., 2020). A number of scholars have begun to examine the potential benefits of inclusive leadership in managing a diverse workforce (Perry et al., 2020; Randel et al., 2018; Shore et al., 2018). Research indicates that inclusive leaders can foster a number of important individual and organizational outcomes including higher levels of job performance, creativity, organizational citizenship behaviors, and team effectiveness (Carmeli et al., 2010; Jin, 2017; Niishi & Mayer, 2009; Panicker, 2018; Randel et al., 2016) and lower levels of turnover (Niishi & Mayer, 2009).

Although scholars acknowledge the important role leaders play in fostering inclusion, there remains little consensus about what inclusive leadership is and how to best measure it (Randel et al., 2016). Historically, scholars viewed inclusive leadership as behaviors that enhance employee voice and foster psychological safety (Carmeli et al., 2010; Nembhard & Edmondson, 2006), but new theory has reconceptualized it as a set of behaviors that foster employees' sense of belonging while allowing them to maintain their uniqueness within the work group (Randel et al., 2018). This new conceptualization is based on the concept of "inclusion," which refers to the inclusion of people with different identities and demographic characteristics (i.e., related to diversity). However, none of the current conceptualizations of inclusive leadership take diverse perspectives into account, particularly in the construction of the current measures being used in

research. Indeed, there are measures of inclusive leadership that have been utilized over the last several years (Zhang et al., 2016), but they do not capture the scope of the construct as identified in the workplace diversity literature. They also do not provide adequate validity or reliability evidence.

The purpose of this dissertation was therefore to: (a) provide clarity about the inclusive leadership construct; (b) develop and validate a theory-based, multidimensional measure of inclusive leadership behavior, the Inclusive Leadership Questionnaire (ILQ); (c) better understand how these ratings are related to employee and organizational outcomes; and (d) to explore how inclusive leadership is perceived by stakeholders with different identities (i.e., on the basis of race and gender). The ultimate goals of this dissertation were to provide a more theoretically grounded foundation for the burgeoning literature on inclusive leadership in organizations and create an empirically validated tool for scholars and leadership development practitioners.

The ILQ, which I developed and gathered validity evidence for in this dissertation, captured the two dimensions of inclusive leadership set forth by Randel et al. (2018; i.e., fostering employees' sense of belonging and valuing their uniqueness), which closely aligns with the predominant definition of inclusion in the organizational diversity literature (Shore et al., 2011). I also integrated a third dimension (i.e., addressing and preventing exclusion), as scholars have proposed the prevention of discrimination, harassment, and other forms of mistreatment as a distinct and important dimension of behavior that leaders must perform to effectively foster inclusive work environments (Perry et al., 2020). To develop and validate my measure, I utilized the traditional six-

phase approach outlined by Hinkin (1998). Phase 1 included item generation using a deductive scale approach (Hinkin, 1998) based on those three theoretical dimensions proposed by Randel et al. (2018) and Perry et al. (2020). Phase 2 included content validation and item reduction. In the content validation step, I provided raters with the definitions of each dimension of inclusive leadership and asked them to sort the items into their respective dimensions. In the item reduction step, I asked subject matter experts (SMEs) to rate the remaining items on relevancy and clarity. In Phase 3, a sample of participants responded to the items of the ILQ, which I then used to conduct a series of exploratory factor analyses to further refine the item pool for each dimension. In Phase 4, a separate sample of participants responded to the remaining items and I utilized those responses to conduct confirmatory factor analyses to confirm the dimensionality of the ILQ. In Phase 5, I assessed the convergent and discriminant validity of the ILQ with six theoretically related measures (e.g., two existing inclusive leadership scales, perceived supervisor support, transformational leadership, allyship, and abusive supervision) and three theoretically unrelated measures (e.g., adventurousness, humor, orderliness). Finally, in Phase 6, I assessed predictive validity of the ILQ by asking participants to rate their supervisors on the ILQ and existing measures of inclusive leadership, and to complete measures for inclusion climate perceptions, job satisfaction, commitment, incivility, and turnover intentions. I then examined the incremental validity of the ILQ in comparison to existing measures of inclusive leadership in predicting those outcomes (see Table 1 for an overview of these phases). This work required seven separate samples with a final total sample of 1,225 individuals. The two outcomes of this dissertation are

(a) a theory-driven measurement tool of inclusive leadership that aligns with current conceptualizations of workplace inclusion, and (b) a better understanding of how employee perceptions of inclusive leadership drive inclusion climate perceptions and other work outcomes.

This work contributes to the field in several important ways. First, it brings clarity to a construct that has been operationalized in multiple ways by examining inclusive leadership more comprehensively. Specifically, it incorporates the needs for not just leader openness, accessibility, and availability (Nembhard & Edmondson, 2006; Carmeli et al., 2010) but also behaviors that will more likely predict positive inclusion climates and prevent exclusion, mistreatment, and discrimination (Perry et al., 2020; Randel et al., 2018). This is particularly timely given the emphasis on inclusion within contemporary organizations as a response to racial tension and gender harassment and discrimination in the United States. Second, the study identifies specific behaviors leaders can enact to be perceived by employees as inclusive. Because existing measures of inclusive leadership have not been comprehensive, they have also not specified behaviors that predict perceptions of inclusion climate. Third, the study produced a rigorously validated tool. Measures based on current definitions of inclusion have utilized the limited definition of inclusive leadership and were also not validated using best practices (i.e., Carmeli et al., 2010; Nembhard & Edmondson, 2006). Fourth, the study is practically relevant, as it can be used by practitioners in various ways (e.g., in employee experience surveys, leadership development). Finally, the measure can serve as an important springboard for future research on inclusive leadership in both academic and applied settings.

Chapter 2: Literature Review

The workforce has grown increasingly diverse over the last several decades, reflected by steadily increasing percentages of women, racial/ethnic minorities, and foreign-born and intergenerational workers (Leslie et al., 2020). Indeed, women now comprise half of the American workforce and hold 40% of managerial positions (Huang et al., 2019) and the current majority, White Americans, is projected to be a numerical minority by 2060 (U.S. Census Bureau, 2012). Starting in the 1990s, many organizations responded to these changes by launching *diversity management* initiatives (e.g., targeted recruiting, diversity training, and mentoring programs), a term popularized by Thomas (1990) to refer to practices aimed at increasing the representation and retention of individuals from diverse backgrounds. Employers asserted that having a diverse workforce may help them drive innovation and creativity, improve decision making, and gain a competitive advantage in an increasingly global business environment (Cox & Blake, 1991), and there is some limited empirical evidence demonstrating diversity promotes improved organizational performance (Jackson & Joshi, 2011). Indeed, a 2018 McKinsey & Co. report found that organizations with diverse boards and executive teams were up to 35 percent more likely to financially outperform their less-diverse competitors.

Diversity is not without its drawbacks, however. Employers have realized it is much easier to create a diverse organization than it is to manage one effectively, particularly as workplaces have become flatter and more reliant on teamwork (Jackson & Joshi, 2011). Empirical evidence confirms this, finding that work group diversity can increase interpersonal conflict and decrease collaboration, leading to decreased

commitment, retention, and satisfaction and higher turnover (Ashikali & Groeneveld, 2015; Jackson & Joshi, 2011). This and other research demonstrates that the effectiveness of diversity management programs alone are limited and largely dependent on the degree to which employees feel included and empowered to fully participate (Ashikali & Groeneveld, 2015; Ng & Stephensen, 2016; Van Knippenberg et al., 2004). For this reason, organizations and scholars have increasingly shifted their focus from "diversity management" (i.e., one focused on solving problems associated with diversity) to "inclusion," which focuses on integrating diverse employees and appreciating and valuing individual differences (Chung et al., 2020; Mor Barak et al., 2016; Nishii, 2013; Roberson, 2006; Shore et al., 2018). Further, inclusion targets both majority and minority group members, encouraging members of all groups to fully participate and bring their "full selves" to work (Chung et al., 2020).

Given the challenges in managing diversity at work, leaders play a critical role in fostering inclusive work environments that in turn help drive organizational objectives. First, as representatives of the organization, leaders have an outsized impact on inclusion. They set and enforce policy and act as role models by setting norms for inclusive behavior (Hollander, 2009). Inclusive leadership has been linked to enhanced inclusion climate perceptions, which in turn can drive higher levels of creativity and performance (Chung et al., 2020). Second, managers can help reduce interpersonal conflict that arises and work to prevent exclusion of individuals from marginalized groups. Ayoko and Konrad (2012) demonstrated that effective leadership can help reduce task and relationship conflicts in diverse teams, which can improve morale and group

performance. Although scholars are increasingly focusing on inclusive leadership (e.g., Perry et al., 2020; Randel et al., 2018), relatively little is known about the specific behaviors inclusive leaders can perform to foster positive inclusion climate perceptions or how to measure them effectively (Randel et al., 2018; Zhang et al., 2016).

The Construct and Measurement of Inclusive Leadership

Research on inclusive leadership is relatively young and has operationalized the construct inconsistently (Randel et al., 2018; Zhang et al., 2016). Nembhard and Edmondson (2006) first introduced "leader inclusiveness" as "words and deeds exhibited by leaders that invite and appreciate other's contributions" (p. 941) and "attempts by leaders to include others in discussions and decisions in which their voices and perspectives might otherwise be absent" (p. 947). Their examination focused on reducing negative effects of high-status differentials within teams (i.e., cross-disciplinary medical teams), and found that inclusive leadership influenced the psychological safety of team members, which in turn influenced their engagement and willingness to speak up. According to their conceptualization, leaders are inclusive to the extent to which they act in ways that foster a psychologically safe environment for all members to voice their perspectives. Carmeli et al. (2010) expanded upon Nembhard and Edmondson's definition, conceptualizing inclusive leadership as openness, accessibility, and availability in interactions with followers. They actually mapped their dimensions of inclusive leadership onto Edmondson's (2004) suggestion that when leaders exhibit openness, availability, and accessibility, they are likely to facilitate the development of psychological safety at work. They found that these behaviors increased employees'

psychological safety, which then facilitated employee involvement in creative work. Although these conceptualizations include behaviors that help shape team members' beliefs that their voices are valued, their focus was limited to fostering employee voice, not a more holistic sense of inclusion as the diversity literature conceptualizes it today. Taking a different approach, Nishii and Mayer (2009) operationalized inclusive leadership as high-quality leader member exchange with all followers (i.e., LMX at the group level) and examined how inclusive leadership might reduce turnover in diverse groups. Within this conceptualization, high differentiation in LMX may indicate some followers have a lower-status position within the group. This suggests leaders who develop quality relationships with all of their followers are seen as more inclusive than those who potentially create divisions within the team. Nishii and Mayer (2009) argued that by building high-quality relationships with followers, leaders signal "their own acceptance of employees of various backgrounds" and can "promote norms about equality and inclusion that will facilitate greater power sharing and improve reciprocal exchanges among group members" (p. 1413). Unlike the other conceptualizations, theirs implies that inclusive leaders must focus on building high-quality relationships with those who might otherwise be left out of such relationships (i.e., marginalized employees; Randel et al., 2018). However, they did not explicitly define what behaviors inclusive leaders should engage in to establish these relationships.

To bring clarity and help establish inclusive leadership as a unique style of leadership that can foster inclusive work climates, Randel et al. (2018) developed a theoretical model of inclusive leadership that includes behaviors that facilitate team

members' sense of belongingness and signal a value for their uniqueness. More recently, Perry et al. (2020) proposed an extension of Randel et al.'s (2018) conceptualization to also include behaviors that actively work to prevent mistreatment. They argue that promoting inclusion is necessary but may not be sufficient for preventing mistreatment of employees who are perceived to be lower status, including marginalized groups (Perry et al., 2020; Pittinsky, 2010). This aligns with Shore et al.'s (2018) updated theoretical model of inclusion that suggests leaders may need to simultaneously promote inclusion and prevent exclusion to foster inclusive work environments and that each may require different types of behaviors.

This new theoretical work is promising because it focuses explicitly on behaviors that enhance employees' perceptions of inclusion, whereas prior operationalizations focused more narrowly on fostering employee voice. As previously mentioned, scholars have increasingly focused on inclusion over the last 20 years because research indicates employees from diverse social and cultural groups are often excluded from opportunities and information networks within organizations (Roberson, 2006; Shore et al., 2011). Most inclusion scholars have defined inclusion as the extent to which individuals can access information and can fully participate in work groups and decision-making processes. In what is now the predominant model of inclusion at work, Shore et al. (2011) defined inclusion as the satisfaction of employees' needs for both belongingness and uniqueness. Their model is based on optimal distinctiveness theory (ODT; Brewer, 1991), which holds that humans have simultaneous competing needs for belongingness (e.g., validation and similarity to others) and uniqueness (e.g., being appreciated for their own

unique characteristics. According to ODT, individuals balance those two needs by finding an optimal level of inclusion in groups to which they belong. To fulfill the human need for belongingness (Baumeister & Leary, 1995), individuals form strong relationships with others (i.e., coworkers, leaders) and seek acceptance into groups (i.e., work teams). Not only are group identities fundamental to one's self-concept (Tajfel & Turner, 1986), but acceptance into groups can help prevent exclusion, particularly if they become highly individuated (i.e., by being a member of a marginalized group). To fulfill the need for uniqueness, individuals need to maintain a differentiated sense of self (Snyder & Fromkin, 1980). According to the theory, if members of groups are perceived as too similar, they risk becoming interchangeable, which puts the need for uniqueness at risk and makes that need more salient. The need for uniqueness relates to the concepts of authenticity (i.e., the degree to which a person's actions are congruent with their beliefs and desires, despite external pressures to conformity; Kernis & Goldman, 2006) and self-determination (i.e., people are authentic when their actions reflect their true or core self, that is, when they are autonomous and self-determining; Deci & Ryan, 1985). Research on ODT in workplace settings has demonstrated that balancing those needs leads to better employee outcomes. For example, Hewlin (2009) found that when employees strive for belonging but suppress their needs for uniqueness, they suffer from emotional exhaustion and report higher turnover intentions. This research supports the argument that organizations should focus on meeting employees' needs for both belongingness and uniqueness in order to foster better work outcomes. Additionally, Chung et al. (2020) confirmed Shore et al.'s (2011) proposed two-factor structure of inclusion as

encompassing both belongingness and uniqueness and confirmed relevant outcomes of inclusion climate (i.e., helping behaviors, creativity, and job performance) proposed by Shore et al. (2011).

Importantly, by mapping their model of inclusive leadership onto this model of inclusion, Randel et al. (2018) brings inclusive leadership into the inclusion literature whereas past operationalizations of inclusive leadership did not. The dimensions of facilitating belongingness and demonstrating a value of uniqueness are types of behaviors that we could assume inclusive leaders should perform. Additionally, behaviors falling under the third dimension proposed by Perry et al. (2020) ensure inclusive leaders are not only promoting a sense of inclusion, but also preventing exclusion. Although this recent theory is promising, there remains no empirical evidence supporting the propositions that these categories of behaviors drive inclusion perceptions at work.

Inclusive Leadership as a Multi-Dimensional and Theoretically Novel Construct

In this section, I formalize a multi-dimensional conceptualization of inclusive leadership that consists of three major dimensions, each with unique sub-dimensions. The first major dimension of inclusive leadership, facilitating belongingness, refers to facilitating employees' needs to form and maintain strong, stable interpersonal relationships. Randel et al. (2018) proposed this dimension includes three types of behaviors: (a) supporting group members, (b) ensuring justice and equality, and (c) sharing decision-making. *Supporting group members* involves creating a comfortable environment, helping team members with their unique needs, and expressing support for them and their opinions. Some specific practices they might institute to achieve these

things include setting up times where they check in with other group members and facilitate conversations where team members share appreciation for colleagues (Randel et al., 2018). *Ensuring justice and equality* involves demonstrating fair treatment of all employees, which communicates to members that they are respected members of the group (Randel et al., 2018). They can achieve this by actively seeking input from all group members before making decisions and considering how decisions could unintentionally create a lack of equity among group members. For example, if a leader suggests that top managers meet at a bar after hours to plan for an important meeting, this could undermine inclusion perceptions for employees who cannot or do not drink, have family care responsibilities after work, or those with a physical disability who rely on limited public transportation (Randel et al., 2018). To help ensure justice and equality, leaders can create systems that aim to reduce bias and offer counter stereotypical information to group members when they notice group members stereotyping their coworkers (Randel et al., 2018; Stone & Colella, 1996). *Shared decision making* refers to sharing power, asking for group-wide participation, and allowing team members to help decide how work is conducted. Inclusion research has established that including employees in decision making is important for fostering belongingness (Mor Barak & Cherin, 1998; Nembhard & Edmondson, 2006; Niishi, 2013; Roberson, 2006). Leaders can do this by asking members to provide input when making major decisions and giving group members opportunities to discuss how to integrate everyone's perspective, or by distributing decision making control over specific aspects of the work so that group members feel included across the group's responsibilities (Randel et al., 2018). Specific

tactics for achieving this could include planning deliberate points in the group's process or workflow where the group comes together to share information and make decisions as a group, and/or using a checklist developed for the purpose of ensuring everyone has had the chance to participate.

The second major dimension, valuing uniqueness, refers to supporting employees' need to maintain a distinctive and differentiated sense of self. Although the inclusion literature primarily emphasizes the importance of belonging (e.g., Mor Barak et al., 1998), individuals' identities and perspectives are considered important factors influencing their perceptions of inclusion as well (Shore et al., 2011). Randel et al. (2018) proposes that inclusive leadership behaviors focused on fostering a sense that uniqueness is valued include: (a) ensuring diverse contributions and (b) helping group members fully contribute. *Ensuring diverse contributions* involves soliciting different points of view, especially from those who may be typically under-represented (Randel et al., 2018). This could help ensure different perspectives are incorporated when the team is considering new approaches to problem solving. At the same time, the inclusive leader must also consider ways to minimize any interpersonal conflict that may arise as they integrate diverse perspectives (Randel et al., 2018). Specific ways leaders can create a welcoming environment that encourages diverse contributions includes forming positive, individualized relationships with each member on their team and recognizing ways that each member is able and willing to contribute (Randel et al., 2018). *Helping group members fully contribute* involves encouraging individuals to contribute who otherwise might not feel their perspectives are welcome (Randel et al., 2018; Roberson, 2006).

Leaders can achieve this by finding ways to involve individuals with different preferences and backgrounds. For example, a leader could ask team members to write down their ideas and then share them one by one, which could encourage people who come from marginalized groups and do not feel as welcome or those who are more introverted to participate (Randel et al., 2018). Inclusive leaders might also find ways to support employees who may need to achieve work tasks and goals in alternative ways due to disabilities (Randel et al., 2018). Additionally, leaders can ensure team members feel they can bring their "full selves" to work (i.e., not downplay or hide differences). While "facilitating belongingness" helps ensure employees' differing perspectives are represented, indicating a value for uniqueness involves making sure employees feel they can fully participate in the group (Randel et al., 2018).

The third major dimension, preventing mistreatment, refers to ensuring compliance with laws that prevent formal discrimination and also confront micro-aggressions to prevent more subtle forms of discrimination (Jones et al., 2016; Perry et al., 2020). Perry et al. (2020) proposed that this dimension includes behaviors such as actively monitoring the level of engagement among participants, identifying factors contributing to exclusion, and working to immediately reduce barriers to engagement. It also entails staying alert to and addressing microaggressions, discriminatory behavior, and exploitation of power differences (Perry et al., 2020).

Inclusive leadership is conceptually related to but distinct from several other styles of leadership (i.e., transformational leadership, authentic leadership, leader-member exchange [LMX], charismatic leadership, ethical leadership, servant leadership,

and participative leadership; Randel et al., 201; Zhang et al., 2016). Specifically, inclusive leadership fosters belongingness and uniqueness and prevents mistreatment in ways that other leader styles do not fully address (see Table 2 for a detailed comparison).

A review of the literature indicates there are seven existing measures of inclusive leadership (Carmeli et al., 2010; Fang et al., 2019; Jin et al., 2017; Nembhard & Edmondson, 2006; Panicker et al., 2018; Ratcliff et al., 2018; Zheng et al., 2017), but they all share two critical limitations: (a) very few of them provided sufficient evidence of validity and reliability, and (b) they are based on various conceptualizations, none of which cover the full scope of inclusive leadership identified in the current diversity literature (see Table 3).

First, none of the measures provided sufficient validity evidence, suggesting they were not validated using best practices in measurement design (see DeVellis, 2003; Hinkin, 1998). Nembhard and Edmondson (2006), who introduced the term "leader inclusiveness," created a three-item measure for their study. Although reliability was adequate ($\alpha = .75$), they did not report any evidence that validity had been assessed. Additionally, their measure was highly contextualized to the medical industry (e.g., "NICU physician leadership encourages nurses to take initiative."). This remains one of the most-used measures by researchers, who typically adapt the measure to refer to a general supervisor (e.g., "My supervisor encourages me to take initiative" or "My supervisor encourages people in my work group to take initiative"). Another most popular measure used by researchers is Carmeli et al.'s (2010) scale, for which they provided some validity evidence. Specifically, they reported having subject matter

experts (SMEs) review items they created for content validation purposes and they performed a factor analysis, but they did not report whether the factor analysis was confirmed with a second sample, nor did they test whether the measure has utility for predicting relevant outcomes over and above the existing measure of inclusive leadership at the time and on which their scale was based (i.e., Nembhard & Edmondson, 2006; both studies predicted psychological safety, so incremental validity could have easily been examined). Jin et al. (2017) reported reliability ($\alpha = .94$) for their 10-item measure, but no validity evidence, stating that their scale was not tested and that they used items similar to those used in previous studies. Zheng et al. (2017) and Panicker et al. (2018) also report reliability ($\alpha = .93$ and .91 respectively) but no validity evidence. Fang et al. (2019) reported some evidence of construct validity (i.e., factor analysis), but they did not provide evidence of convergent and discriminant validity, nor did they test the utility of the scale to predict relevant outcomes above and beyond existing scales. Ratcliff et al. (2018) reported how their items were developed, reviewed by SMEs, further refined, and then reviewed by external researchers for content validation, but did not test their item pool further. Their paper was a research report from the Army Research Center, and they outlined their next steps to validate their survey in a "Future Survey Development Work" section. Importantly, the scale is highly contextualized to a military context (i.e., one dimension is called "Integration into the unit"; sample item: "Pairs new unit members with experienced Soldiers when performing tasks").

Second, each of the existing measures conceptualize inclusive leadership a bit differently and none cover the full scope of inclusive leadership identified in the current

diversity literature. Nembhard and Edmondson (2006) introduced "leader inclusiveness" to the literature and operationalized it as inviting and appreciating contributions. Carmeli et al.'s (2010) measure built off of Nembhard and Edmondson (2006), conceptualizing it as involving openness, availability, and accessibility. Although the leader behaviors included in these measures encourage individuals to share their perspectives, they do not directly assess leader behaviors that facilitate employees' sense of belongingness, signal a value for uniqueness, and prevent mistreatment. Fang et al. (2019) stated they based their measure on Carmeli et al. (2010), however, they then report their factor analysis identified three factors (i.e., the leaders' encouragement and recognition of employees, the leaders' respect and fair treatment of employees, and leaders' rational understanding and tolerance of employees' failures) that are different from Carmeli et. al's three factors [openness, availability, and accessibility]). Fang et al.'s "respect and fair treatment" behaviors may cover some elements of "preventing exclusion," but are not explicitly focused on doing so (i.e., a leader might personally treat followers fairly, but may not step in if they notice co-workers mistreating each other). Again, they did not report their full list of items, but because their measure was not grounded in theory, they likely do not cover the full scope of behaviors that would foster an inclusive workplace. Similarly, Jin et al. (2017) stated they created items similar to those used in past studies (i.e., Carmeli et al., 2010). However, they then state "leadership involves leader supportiveness, sensitivity to follower needs, open communication, and fairness," which do not align with Carmeli et al.'s (2010) primary dimensions. Panicker et al., (2018) defined inclusive leadership as acknowledging the contribution of subordinates and involving them in

various activities, citing Nembhard and Edmondson (2006), but these authors generated items for their measure related to fair treatment (e.g., "I am treated fairly and with dignity"). However, they do not ground their measure in theory and do not report their full list of items (nor do they perform factor analysis), so it is unclear what dimensions they propose inclusive leadership encompasses. The extension of these measures into the domain of fairness was promising because inclusion does require a sense that one will be treated fairly (Ferdman et al., 2014; Mor Barak et al., 1998; Perry et al., 2020). However, just because one is treated fairly and with dignity by their supervisor does not mean they will be treated fairly by coworkers, and inclusive leaders must do that to truly foster inclusive work climates. Zheng et al. (2017) created their items based on Pless and Maak's (2004) description of inclusion, with items that cover appreciation, respect, open communication, participative decision making, moral reasoning, and a cooperative working style. This measure includes behaviors that may be expected to foster a sense that one's uniqueness is valued, but they do not target individual employees' belongingness needs or the prevention of mistreatment. Finally, Ratcliff et al. (2018) conceptualized inclusive leadership as encompassing five dimensions: fair treatment, openness to differences, integration into the unit, leveraging unique perspectives and expertise, and shared understanding in communications. They grounded their measure in Randel et al.'s (2018) theory of inclusive leadership as behaviors that facilitate belongingness and uniqueness, but added dimensions unique to the military context and based on Brown et al.'s (2018) five-dimension model of an inclusive military climate.

This measure is quite specific to a military context and also does not include behaviors aimed at preventing mistreatment of followers.

For the above reasons, it is no surprise that in a paper focused on the measurement of inclusive leadership, Zhang et al. (2016) called for future research because the measurement of this construct is weak. Effective measurement is key to understanding the impact of leader behaviors on employee perceptions of inclusion and associated outcomes. It is also imperative for developing theoretically and practically relevant interventions.

Hypothesis Development

The purpose of this study was to develop a measure of inclusive leadership behaviors using dimensions delineated by Randel et al. (2018) and Perry et al. (2020). This is important because inclusive leadership is associated with individual, team, and organizational outcomes. Specifically, inclusive leadership is associated with individual outcomes such as increased job satisfaction (Findler et al., 2007), organizational commitment (Cho & Mor Barak, 2008; Findler, & Wind, 2001; Findler et al., 2007; Nembhard & Edmondson, 2006), job performance (Cho & Mor Barak, 2008; Chung et al., 2020), creativity and innovation (Carmeli et al., 2010; Choi, Tran, & Park, 2015; Chung et al., 2020), work engagement (Choi, Tran, & Park, 2015; Nembhard & Edmondson, 2006), inclusion climate (Chung et al., 2020), psychological safety (Carmeli et al., 2010; Nembhard & Edmondson, 2006), well-being (Findler et al., 2007), and helping behaviors (Chung et al., 2020; Panicker, 2018; Randel et al., 2016). It is also related to lower levels of turnover intentions (Nishii & Mayer, 2009). Team level

outcomes of inclusive leadership include participation in quality improvement efforts (Howard et al., 2012; Nembhard & Edmondson, 2006), team psychological safety climate (Hirak et al., 2012), team identification (Mitchell et al., 2015), lower perceptions of status differences (Niishi & Mayer, 2009), higher team performance (Hirak et al., 2012; Mitchell et al., 2015), and lower turnover (Niishi & Mayer, 2009). Organizational outcomes of inclusive leadership include perceptions of fairness, and management of diversity (Ryan, 2006; Devecchi & Nevin, 2010; Granados & Kruse, 2011).

Given that I proposed to use a deductive scale development approach to generate items (i.e., SMEs generated items based on inclusive leadership dimensions theorized by Randel et al. [2018] and Perry et al. [2020]) and further proposed trimming items in an exploratory factor analysis sample, I hypothesized the following for the confirmatory factor analysis sample:

> *Hypothesis 1:* A three-factor model of the ILQ, in which the items are set to load onto their respective factors ("facilitating belongingness" [A], "valuing uniqueness" [B], and "preventing exclusion" [C]), will fit significantly better than a one-factor model *(H1a)*, a two-factor model with A and B together *(H1b)*, a two-factor model with A and C together (*H1c*), and a two-factor model with B and C together *(H1d)*.
>
> *Research Question 1:* Are second-order factor loadings sufficiently large to support the idea of a common second-order factor ("inclusive leadership")? See Figure 1 for my hypothesized hierarchical model.

To establish the nomological network of the ILQ, I assessed the convergent and discriminant validity of this measure with other theoretically relevant and disparate constructs. In particular, to establish convergent validity, I examined the ILQ's relation to two existing scales of inclusive leadership (Carmeli et al., 2010, Nembhard & Edmondson, 2006) along with four theoretically related measures: perceived supervisor support, transformational leadership, allyship behaviors, and abusive supervision.

Although existing measures of inclusive leadership (i.e., Carmeli et al., 2010; Nembhard & Edmondson, 2006) include a more limited set of inclusive leadership behaviors, I anticipated they would nevertheless positively correlate with my measure because they emphasize supportive leader behaviors such as openness, availability, and accessibility.

Perceived supervisor support is conceptualized as employees' perceptions that their supervisor values their contributions and cares about their well-being (Eisenberger et al., 1986). As such, it should positively relate to a set of leadership behaviors focused on ensuring employees feel they belong, can be themselves at work, and do not feel mistreated.

Transformational leadership is a leadership style in which leaders encourage, inspire and motivate employees to innovate and create change and includes four categories of behavior: inspirational motivation, intellectual stimulation, idealized influence, and individual consideration (Bass & Avolio, 1993). Individual consideration is defined as the leader's ability to build unique relationships with each follower, understanding that each person may have unique needs and abilities (Bass & Riggio,

2006). This aligns with inclusion theory (Shore et al., 2011), which suggests that leaders who treat employees as unique individuals may help them feel valued for their uniqueness, a critical part of inclusion. Leaders who pay attention to individual followers' needs would likely also notice opportunities to prevent mistreatment and exclusion (Perry et al., 2020). Intellectual stimulation is defined as the leader's ability to help followers consider problems from new perspectives and to stimulate critical and creative thinking (Bass & Avolio, 1997). Leaders can do this by seeking out differing perspectives and encouraging followers to look at problems from different angles (Bass & Bass, 2008). Expressing an openness for diverse viewpoints may lead followers to feel their uniqueness is valued and thus increase perceptions of inclusion (Brimhall & Palinkas, 2020; Nembhard & Edmondson, 2006). Inspirational motivation is defined as the leader's ability to share a compelling vision for the future (Bass & Avolio, 1997). Idealized influence refers to the leader's ability to serve as a role model for followers and increase followers' identification with, trust in them, and desire to emulate them (Bass & Avolio, 2008). When employees perceive they are similar to their leader, they may feel increased belonging, a critical component of inclusion (Brewer, 1991; Shore et al., 2011). These two components together (sharing an inspiring vision for the future and increasing a sense of commonality with the leader) would likely create a shared sense of purpose and thereby foster a sense of belongingness to the organization (Bass & Bass, 2008; Brimhall & Palinkas, 2020). Indeed, research has found that all four components of transformational leadership positively relate to inclusion climate (Brimhall, 2019;

Brimhall & Palinkas, 2020; Chrobot-Mason et al., 2014). I therefore anticipated inclusive leadership would positively relate to all four dimensions of transformational leadership.

Allyship refers to behaviors by individuals who work to end discrimination for oppressed groups (Washington & Evans, 1991). Allies may act in ways that demonstrate support (i.e., providing comfort and tangible resources for individuals with stigmatized identities, attending educational events geared toward minority groups, or showing acceptance and positivity when receiving disclosures of invisible stigmas from co-workers; Sabat et al., 2013) or advocacy (i.e., more public demonstrations of support for issues affecting stigmatized individuals such as championing diversity initiatives, requesting additional resources or supportive policies for minority employees, or directly confronting discrimination or harassment; Sabat et al., 2013). It is likely that when employees view their leaders as inclusive (i.e., as valuing individuals' unique identities and working to prevent exclusion), they would also view them as exhibiting behaviors indicative of allyship. I therefore anticipated inclusive leadership would positively relate to both support and advocacy dimensions of allyship behaviors.

Finally, abusive supervision refers to subordinates' perceptions of supervisors' display of hostile verbal and nonverbal behaviors, excluding physical contact (e.g., public criticism, loud and angry tantrums, rudeness, inconsiderate actions, and coercion; Tepper, 2000). Because abusive supervision is a form of mistreatment (Hershkovis, 2011), it is unlikely that followers would perceive supervisors as fostering belongingness, valuing them as individuals, and preventing exclusion when they exhibit abusive supervisor

behaviors. I therefore expected inclusive leadership to negatively relate to abusive supervision.

To establish the discriminant validity of the ILQ, I examined its relation to measures of adventurousness, humor, and orderliness. Adventurousness has been defined as having a wide variety of interests and being imaginative and optimistic and is a facet of openness to experience on the five-factor model (FFM) of personality (Costa & McCrae, 1992). Humor has been defined as the ability to make others smile or laugh (Peterson & Seligman, 2004). Orderliness has been defined as being organized, precise, thorough, and methodical, and is a facet of conscientiousness on the FFM (Costa & McCrae, 1992). There is currently no literature that suggests inclusive leadership should relate to these three personality constructs. As such, I hypothesized the following:

> *Hypothesis 2:* The Inclusive Leadership Questionnaire (ILQ) will demonstrate convergent validity (moderately correlate) with existing measures of inclusive leadership (*H2a*), perceived supervisor support (*H2b*), transformational leadership *(H2c),* allyship (*H2d*), and abusive supervision (*H2e*).
>
> *Hypothesis 3*: The ILQ will demonstrate discriminant validity (low correlations) with measures of adventurousness (*H3a*), humor (*H3b*), and orderliness (*H3c*).

Beyond construct validity, it is important to establish the utility of my new measure. Specifically, I investigated the ability of the ILQ to predict relevant outcomes (i.e., inclusion climate perceptions, job attitudes, incivility, and turnover intentions) and that it does so beyond existing scales of inclusive leadership (i.e., Carmeli et al., 2010; Nembhard & Edmondson, 2006).

First, I examined the predictive validity of the ILQ on inclusion climate perceptions, job satisfaction, affective commitment, incivility, and turnover intentions. Inclusion climate perceptions refer to the degree to which an employee perceives their needs for belongingness (i.e., feeling like they are treated as an insider) and uniqueness (i.e., feeling like they are encouraged to retain their uniqueness) are satisfied by the work group and the organization (Shore et al., 2011). This definition is consistent with optimal distinctiveness theory (Brewer, 1991), which posits that individuals seek balance between two opposing needs: belonging to groups of similar others and maintaining a sense of uniqueness and autonomy. In work contexts, if these needs are not met (i.e., if one's sense of belongingness or the perceived value of uniqueness is low), inclusion perceptions will be low (Shore et al., 2011). Importantly, these perceptions are shaped not only by how individuals are treated but also how they perceive their coworkers are treated (Perry et al., 2020). When leaders act in ways that foster employees' sense of belongingness and signal that they value their uniqueness, employees' perceptions of inclusion climate should be high.

Inclusive leadership behaviors should also predict higher levels of job satisfaction and affective commitment, and lower levels of turnover and mistreatment at work (i.e., incivility). Job satisfaction is defined as the sense of enjoyment or fulfillment one gets from their job (Dalal, 2013). There are several facets of job satisfaction (i.e., satisfaction with one's pay, opportunities, supervisor, coworker, and nature of the work), and person-environment fit theories suggest that if an employee perceives a discrepancy between what they need/want and what the job is able to supply, they will be dissatisfied with

their job (Dalal, 2013). It is therefore likely that if employees do not perceive their leaders are acting in ways that meet their fundamental needs for belongingness and relatedness, they will feel dissatisfied. Research indicates that when employees feel their organizations have effectively managed diversity and fostered a climate of inclusion, they report higher levels of job satisfaction (Acquavita et al., 2009; Mor Barak et al., 2016).

Affective commitment refers to an emotional bond that an employee feels with the organization (Allen & Meyer, 1990). Inclusive leadership will likely lead to higher levels of affective commitment among followers because enhancing their sense of belongingness will help employees identify with the organization (Ashikali & Groeneveld, 2015). Research indicates that when employees feel their organizations have effectively managed diversity and fostered a climate of inclusion, they report higher levels of affective commitment (Mor Barak et al., 2016).

Incivility is defined as low-intensity deviant behavior (e.g., rudeness; discourteousness) that violates workplace norms for mutual respect (Pearson & Porath, 2005). Research has found that positive leader behaviors (i.e., those associated with ethical and charismatic leadership) can lead to lower levels of incivility by enhancing employees' perceptions of norms for respect (Walsh et al., 2017). It is likely that leaders who work to foster belongingness, value uniqueness, and prevent mistreatment would also enhance employee perceptions of norms for respect and thereby reduce the likelihood of incivility occurring in their work environments. Additionally, multiple lines of research have found that inclusive leadership predicts higher levels of helping

behaviors (Ashikali & Groeneveld, 2015; Chung et al., 2020), and helping behaviors negatively relate to incivility (Pearson & Porath, 2005).

Turnover intentions refer to the degree to which employees plan to leave their jobs. Griffeth, Hom, and Gaertner's (2000) meta-analytic findings demonstrated turnover intentions to be the strongest single predictor of actual voluntary turnover (corrected mean r = .45). If an employee perceives they belong and are valued as an individual at work, they would be less likely to leave their job. Past research indicates that inclusive leadership negatively relates to turnover intentions (Nishii & Mayer, 2009). As such, I hypothesized the following:

> *Hypothesis 4:* The ILQ will predict inclusion climate perceptions (*H4a*), job satisfaction (*H4b*), affective commitment (*H4c*), incivility (*H4d*), and turnover intentions (*H4e*).

In terms of incremental validity, the ILQ should more strongly predict inclusion climate compared to existing measures of inclusive leadership (i.e., Carmeli et al., 2010; Nembhard & Edmondson, 1996) because those measures are based on narrower definitions of inclusive leadership that do not consider the management of a diverse workforce. Specifically, these measures define inclusive leadership in terms of a leader's "openness and availability" and do not consider behaviors that foster a sense of belonging or acceptance of uniqueness at work. Additionally, these existing measures have not been empirically validated using established best practices. As such, I hypothesized the following:

Hypothesis 5: The ILQ will be a stronger predictor of inclusion climate perceptions than existing measures of inclusive leadership (i.e., Carmeli et al., 2010; *H5a*); Nembhard & Edmondson, 2006; *H5b*).

Finally, my measure may better predict organizational outcomes (i.e., job satisfaction, affective commitment, incivility, and turnover intentions) than existing measures of inclusive leadership. As stated previously, prior conceptualizations of inclusive leadership have negatively predicted turnover (Nishii & Mayer, 2009), and positively related to job satisfaction (Findler et al., 2007) and commitment (Choi et al., 2015; Findler et al., 2007). However, those prior conceptualizations, which emphasized employee voice, did not include preventing mistreatment and fostering a sense of belongingness, which would likely further enhance these outcomes, though there is currently not enough literature to strongly support such a hypothesis. Therefore, I will explore the following research question.

Research Question 2: Will the ILQ be a stronger predictor of job satisfaction (*RQ2a*), commitment (*RQ2b)*, incivility (*RQ2c)*, and turnover intentions (*RQ2d)* than existing measures of inclusive leadership?

Finally, employees who hold under-represented or marginalized identities based on personal characteristics (e.g., race, gender, age, religion, ability, sexual orientation, gender identity) would likely report lower levels of perceived inclusion climate than majority group employees (e.g., Whites, men), which could strengthen the impact of inclusive leadership behaviors on their workplace experiences compared to majority

group employees. Indeed, perceptions of inclusion climate often differ among employees of different demographic backgrounds, with ethnic minorities and women perceiving lower levels of inclusion at work (Kossek & Zonia, 1993; Mor Barak et al., 1998). Specifically, Mor Barak et al. (1998) found that White men had the most positive perceptions and women of color had the least positive inclusion climate perceptions. Other research indicates that women and members of racial/ethnic minorities are more likely to feel excluded at work and that exclusion is linked to job dissatisfaction and lower sense of well-being (Mor Barak & Levin, 2002). Other research demonstrated that the positive effect of work group inclusion on health is stronger when individuals are in the numerical minority in their work group in terms of race or gender (Ehrhart et al., 2014).

Alternatively, members of majority groups (e.g., Whites, men) may interpret and react more strongly to inclusive leader behaviors than employees who hold under-represented or marginalized identities. For example, Whites and men may perceive some inclusive leadership behaviors as excluding them (e.g., asking women and minority employees to share their thoughts and opinions; advocating for policies that benefit minority group members), leading to lower ratings on inclusion climate among majority group members. For example, research indicates that members of high-status groups (i.e., Whites, men) perceive pro-diversity messages from their organizations as threatening (Dover et al., 2016). Additionally, using a multi-cultural approach to diversity management (which seeks to identify, articulate, and utilize differences) can be met with

higher levels of backlash than a "colorblind" approach, which downplays group differences (Leslie et al., 2019). I therefore explored the following research question:

> *Research Question 3*: How will identity characteristics (i.e., race and gender) impact the interpretation and reactions to inclusive leadership behaviors (i.e., through their responses to inclusion climate perceptions)?

Chapter 3: Study Method and Results

The goal of this dissertation was to develop and provide initial evidence for reliability and validity of this measure of inclusive leadership based on theory. To do this, I followed study design recommendations from Hinkin (1998) and DeVellis (2003). Specifically, my study included six phases: (a) item generation, (b) content validation/item reduction, (c) exploratory factor analysis, (d) confirmatory factor analysis, (e) convergent/discriminant validity, and (f) predictive validity. The goal was to develop a scale with three dimensions and approximately four to six items per dimension, in line with best practices for measure development (Hinkin, 1998).

Phase 1: Item Generation

Because the goal of the measure development was to obtain 4-6 items per dimension, and because approximately 50% of the originally generated items are expected to be retained (Hinkin, 1998), I recruited 12 subject matter experts and asked each to develop three to five items per dimension (i.e., facilitating belongingness, valuing uniqueness, preventing mistreatment, and a general category for additional inclusive leadership items), or potentially 240 items.

Sample. Twelve SMEs in the area of workplace diversity, inclusion, and leadership, 10 of whom have experience with managing a traditionally marginalized identity in the workplace (i.e., on the basis of gender, race, ethnicity, and/or sexual orientation), were asked to assist with the generation of items for this study. This sample included one faculty member in Industrial-Organizational (I-O) Psychology, one

doctoral-level practitioner in the area of I-O psychology, one practitioner in workplace justice, equity, diversity, and inclusion, six graduate students in I-O Psychology, and three practitioners in research-related roles. In terms of expertise, this sample was in line with recommendations by Hinkin (1998).

Design. Because the theoretical foundation for the construct of inclusive leadership is clearly defined, I used a deductive scale development approach (Hinkin, 1998). I provided SMEs with the following definitions for each of the three inclusive leadership dimensions outlined by Randel and colleagues (2018) and Perry and colleagues (2020):

> *Facilitating belongingness:* Facilitating employees' needs to form and maintain strong, stable interpersonal relationships. This includes behaviors involving ensuring justice and equity, sharing decision making, and supporting individuals as group members.
>
> *Valuing Uniqueness:* Supporting employees' need to maintain a distinctive and differentiated sense of self. This includes behaviors involving encouraging diverse contributions and helping group members fully contribute.
>
> *Addressing and preventing exclusion:* Ensuring compliance with laws to prevent formal discrimination as well as directly confronting microaggressions toward low-status team members to prevent subtle discrimination.

I instructed SMEs to write items that represent each dimension in reference to behaviors of leaders (i.e., defined as people in organizational contexts who oversee other employees) and provided one example item per dimension. I also provided them with specific instructions that align with a list of best practices in item generation, outlined by Hinkin (1998) and others (DeVellis, 2003; Spector, 1992; see Appendix A). First, items should be as simple and concise as possible, as unnecessarily lengthy items can reduce

clarity. Specifically, items should be made up of short words in order to remain at a 5th-7th grade reading difficulty level, which is considered appropriate for a measure's use in the general population (DeVellis, 2003; Fry, 1977). Additionally, items should not contain jargon, expressions, or colloquialisms, including ones that may be time-constrained (i.e., "My leader supports the MeToo movement"; Spector, 1992).

Second, items should not contain multiple negatives (i.e., "My leader does not stop others from discriminating against me") because such items may confuse respondents. Third, items should contain only one inclusive leader behavior; double-barreled items may represent two separate constructs and confuse respondents. Fourth, items should remain consistent in perspective (i.e., not mix affective and behavioral responses). SMEs were therefore reminded to write items focused on behaviors leaders can enact, not statements that reflect affective responses to or attitudes toward leaders (Harrison & McLaughlin, 1993). Fifth, I reminded SMEs to avoid writing leading questions, which may bias responses, and to instead aim for items that are generally neutral in valence so that participants can more freely endorse them.

Sixth, I instructed SMEs that they may write redundant items because it can be helpful in establishing internal consistency reliability (DeVellis, 2003). In order for redundant items to be useful, the redundancy should be limited to content (i.e., not also redundant in sentence structure or terminology), so as to capture enough variance. Seventh, I instructed SMEs to use only positively worded items. Although reverse-scoring items remains a technique for reducing response-set bias (i.e., acquiescence, affirmation, or agreement bias; Price & Mueller, 1986) and inattentive responding, such

items can confuse respondents and lead to negative psychometric outcomes (DeVellis, 2003; Harrison & McLaughlin, 1991; Hinkin, 1998). Instead, I utilized other techniques (i.e., instructed response and bogus items) to assess inattentive responding.

Each SME developed items without having access to any other SMEs' developed pool of items. Doing so prevented SMEs from being influenced by others' items. I anticipated some level of redundancy across SME item pools, and therefore oversampled in terms of the total number of items generated. This allowed for the removal of overly redundant items or items that did not meet the best practices and guidelines for item generation I provided to the SMEs.

This process resulted in a total of 231 items generated by SMEs. Two SMEs (a faculty member in Applied Psychology and myself) then screened the items to ensure they followed the guidelines listed above (e.g., double-barreled items were modified or removed) and were high quality and not overly redundant. This resulted in a pool of 82 items.

Phase 2: Content Validation/Item Reduction

After SMEs drafted an initial pool of items, I examined the content validity of this set of items (Hinkin, 1998). Specifically, I examined content adequacy with an item sorting task, followed by a review for clarity and relevance.

Content Adequacy Sample. I recruited 96 working adult students at a university in the Northwest United States by working with instructors to offer extra credit to students who participated in the study. Because this is a cognitive task that does not require an understanding of the phenomena under investigation, it is appropriate to use a

sample of students (Anderson & Gerbing, 1991; Schriesheim et al., Powers, Scandura, Gardiner, & Lankau, 1993). Because I was interested in developing a measure that can be used within organizations, I did, however, limit the sample to employed adults who worked at least 20 hours per week under the direction of a supervisor. To do this, I embedded inclusion items into the demographic questionnaire at the end of the survey. If respondents indicated they were not currently working, or were not working 20+ hours per week under the supervision of a manager, I removed their cases from analyses. To help ensure high-quality data, I also included instructed response items to screen for inattentive responding (e.g., "Please match this item with Facilitating Belongingness"), following established best practices by Meade and Craig (2012). Fifty-one respondents did not meet inclusion criteria or did not pass at least one of the attention checks, resulting in a final sample of 45 participants (M_{Age} = 26.15, SD_{Age} = 6.61; 60% Women; 69% White, 18% Asian, 13% Hispanic/Latinx, and 7% Black/African American; 62% employed full-time). Per Hinkin (1998), this is an appropriate sample size for this task (Anderson & Gerbing, 1991; Schriesheim et al, 1993).

Design and Results. I presented participants with definitions of the dimensions and asked them to match individual items with each of the dimensions. I also included an "unclassified" category so that participants could choose to identify items as not representing any dimension. Then, in line with Hinkin's (1998) recommendations, I evaluated items for content validity with Anderson and Gerbing's (1991) proportion of substantive agreement (defined as the number of respondents who classified an item to its intended construct over the total number of respondents). I retained items with a

proportion of substantive agreement of at least .51 (indicating more than half of the participants correctly classified the item). This resulted in 57 items progressing to the next stage.

Item Reduction Sample. To further evaluate content validity, I then recruited 12 SMEs in the areas of workplace diversity, inclusion, and leadership who also hold one or more traditionally marginalized identities (i.e., race, ethnicity, gender, sexual orientation, etc.) to assist with screening items for relevance and clarity (DeVellis, 2003). These SMEs were different individuals from those who participated in the initial item generation.

Design and Results. SMEs completed a survey in which they rated the items in terms of their "relevance" and "clarity." Specifically, SMEs rated each item on a unipolar scale from 1 ("not at all") to 5 ("extremely") with regard to their clarity and relevance to the dimension of inclusive leadership they represent. As recommended by DeVellis (2003) another SME and I then examined items that were rated as relatively lower on relevance and clarity and removed such items if deemed appropriate. This process resulted in 41 items with average ratings of at least "4" ("strongly") being retained.

Phase 3: Exploratory Factor Analysis (Initial Item Reduction and Internal Consistency Assessment)

I then further refined the item pool with a series of exploratory factor analyses. My goal was to obtain between 4-6 items per factor per recommendations by Hinkin (1998). Because I planned to use the same sampling strategy for both exploratory and

confirmatory factor analysis stages, I collected one large sample (n = 3,741) and randomly split the data into two samples.

Sample. To analyze the factor structure of the generated items, I sought a sample of 275 working adults in the US. Scholars have provided some rules of thumb in terms of generally acceptable sample sizes required to obtain accurate factor solutions, with recommendations for item-to-response ratios ranging from 1:4 (Rummel, 1970) to more than 1:10 (Schwab, 1980). Other research has found, however, that a sample size of 150 should be sufficient to obtain accurate solutions given reasonably strong item intercorrelations (Guadagnoli & Velicer, 1988). Because I retained 41 items, I sampled 275 participants with an item-to-response ratio of approximately 1:7 for the exploratory factor analysis.

To recruit participants who met my inclusion criteria of being located in the US and were employed full-time, I utilized Amazon MTurk, a crowdsourced work platform on which individuals can receive compensation for completing tasks' requested by others. Because MTurk's platform is open to workers from a variety of backgrounds, job statuses, organizations, and industries, it is an appropriate sampling source because this measure aims to generalize across organizations and industries.

With regards to data quality, research has demonstrated that data obtained from MTurk studies are at least as reliable as those obtained via traditional methods (Ran, Liu, Marchiondo, & Huang, 2015; Fleischer, Mead, & Huang, 2015). Indeed, a recent meta-analysis based on 90 independent samples found that data collected on online panels including MTurk had similar psychometric properties and produced criterion validities

that converged with those of conventionally sourced data (Walter et al., 2018). Some research indicates that there could be a danger in using crowdsourced study data without restrictions. I therefore utilized CloudResearch, an online tool that allows researchers to tailor their sample recruitment strategy on MTurk, to improve data quality. Specifically, based on recommendations by Feitosa et al. (2014), who found that crowdsourced data were similar to traditionally collected data only when they were restricted to IP addresses from English-speaking countries, I restricted the current sample to participants within the US. I also utilized settings to restrict the study to only those participants who had at least a 90% approval rating on MTurk and had completed 50 or more previous MTurk surveys, which are suggested best practices for ensuring quality data (Bartel-Sheehan & Pittman, 2016; Peer, Vosgerau, & Acquisti, 2014). Hillygus, Jackson, and Young (2014) found less bias in frequent responders to online panels than in infrequent survey responders. Although some research indicates that MTurk can produce representative samples and that MTurk participants are often more demographically diverse than those of typical psychological studies such as American college samples and standard Internet samples (Buhrmeister et al., 2011; Walter et al., 2018), our own past research has found MTurk samples to be overwhelmingly (i.e., 80% +) White. Therefore, in an attempt to ensure a diverse sample, I created two separate MTurk HITs, one recruiting white participants and one recruiting non-White participants.

I again utilized screening techniques recommended by Smith and colleagues (2015) to ensure participants were employed adults working at least 20 hours per week under the direction of a supervisor. I again used instructed response items (e.g., “Please

select fairly often for this one."), as recommended by Meade and Craig (2012). Participants who did not correctly answer all three of the instructed response items were redirected to an end-of-survey message on CloudResearch's platform, which captured their IP address and did not allow them to reattempt to take the survey.

Because I planned to use the same sampling strategy for both exploratory and confirmatory factor analysis stages, I recruited a total sample of 3741 working adults (n = 3741) which included those who did not meet inclusion criteria. I removed individuals who did not meet inclusion criteria or failed attention checks, resulting in a final sample of 548 participants (M_{Age} = 41.41; SD_{Age} = 12.03; 66% Women; 55% White; 26% Black/African American; 13% Hispanic/Latinx; 9% Asian; 93% employed full-time). I then randomly split the data for exploratory (n = 275) and confirmatory (n = 273) factor analyses. For the EFA sample, the average age was 42.11 (SD_{Age} = 11.81), 66% were women; 41% White; 24% Black/African American; 14% Hispanic/Latinx; 9% Asian; 92% employed full-time.

Design: I instructed participants to provide the frequency with which they have observed their supervisor at work enact each behavior on a 5-point unipolar frequency scale (1 = "Never, 5 = "Frequently, if not always"). The instructions included a note advising them to, if they have not directly observed a behavior, please rate how frequently they believe or assume they would behave in that way, based on what they know of their leadership style, personal characteristics, or interpersonal relationships. I reminded participants that their responses would be kept completely confidential and requested that they answer honestly. I once again embedded multiple instructed-response

and bogus items into the survey to help ensure data quality (Meade & Craig, 2012). After responding to the scale items, participants provided additional demographic information.

Results. I used the *lavaan* package in R to perform exploratory factor analyses. I first examined all items and found they met assumptions of normality (standardized skewness ≤ |1.06| and standardized kurtosis ≤ |1.36|; Hair et al., 2009; Hair, Black, Babin, & Anderson, 2009). I then examined the inter-item correlations of the variables for each factor and all items were well above the recommendation of $r < .40$ (in all cases, $r > 0.50$; Kim & Mueller, 1978). I then conducted an initial principal components analysis and examined the Bartlett's (1950) test of sphericity ($p < .001$), and Kaiser-Meyer-Olkin measure of sampling adequacy (KMO = .98; acceptability is > .60; Tabachnick & Fidell, 2012), which confirmed the factorability of the data. To determine how many factors to extract, I drew upon my theoretical model of three factors. Although only two factors demonstrated eigenvalues over 1 (a rule of thumb for empirically deciding how many factors to extract; Tabachnick & Fidell, 2012), a third factor was close to meeting that threshold (.94). Additionally, because I had a strong theoretical and methodological rationale and results at each step produced interpretable solutions, I continued to extract three factors at each step.

I used a common factor analysis with maximum-likelihood extraction because I was interested in the latent constructs for each factor (Hinkin, 1998). Additionally, I used Promax rotation, an oblique rotation, because I expected the factors to be correlated (DeVellis, 2003; Fabrigar et al., 1999; Hendrickson & White, 1964). With each iteration, I retained items that demonstrated factor loadings greater than .40 with no cross-loadings

greater than .30 (Ford et al., 1986; Hinkin, 1998). This iterative process resulted in 24 items being retained. For the sake of parsimony, I removed an additional 13 items that were duplicate in content to another item within the same factor. This resulted in a final set of items that demonstrated a clear factor structure: three items belonging to the "facilitating belongingness" factor, six items to the "supporting uniqueness" factor, and seven items to the "preventing mistreatment" factor, with minimal cross-loadings (see Table 4). Reliabilities surpassed acceptable levels for the full scale (α = .97) and for each factor (Nunally, 1978): facilitating belongingness (α = .95), valuing uniqueness (α = .91), preventing mistreatment (α = .96). Correlations between factors ranged from .72–.77. The percentage of total item variance that is explained was 75%, well above the minimum acceptable target of 60% (Hinkin, 1998).

Phase 4: Confirmatory Factor Analysis

After developing a final set of 14 items with three to seven items per factor, I wanted to confirm the factor structure in a separate sample by testing a series of models with confirmatory factor analysis (CFA). CFA allows the researcher to assess the goodness of fit of the hypothesized model and allows for the comparison of the hypothesized, multidimensional model to a series of alternative models in which dimensions are combined (Hinkin, 1998). Using a separate sample also allows for replication of the factor structure identified in the exploratory factor analysis sample (Hinkin, 1998). Further, CFA allows for the testing of a second-order factor model, providing evidence that a summated score could be used for the three subscales.

Sample. After randomly splitting the final sample of 548 respondents who met inclusion criteria and passed all attention checks, a sample size of 273 participants were utilized for confirmatory factor analysis (M_{Age} = 40.70, SD_{Age} = 12.22; 65% Women; 52% White; 29% Black/African American; 11% Hispanic/Latinx; 9% Asian; 94% employed full-time). This sample size aligns with Hinkin (1998), who recommends a minimum sample size of 200 for CFA. Additionally, this sample is demographically similar to the prior exploratory factor analysis sample.

Results. I used the *lavaan* package in R to perform confirmatory factor analyses. All items met acceptable levels of normality (all cases standardized skewness ($\leq |1.24|$) and kurtosis ($\leq |1.22|$; Kline, 2010). One case was missing one response. Once again the full scale (α = .97) and individual factors ($\alpha_{belongingness}$ = .95, $\alpha_{uniqueness}$ = .92, $\alpha_{mistreatment}$ = .96) demonstrated strong reliability.

I estimated all models with the maximum likelihood estimation technique (as recommended by Kline, 2010). First, to test my hypothesized three-factor model, I specified a model in which each item was set to load onto its respective hypothesized latent factor. Because the chi-square statistic is sensitive to sample size (Hinkin, 1998), I assessed goodness of fit with indices including CFI, RMSEA, SRMR, and TLI to determine the acceptability of the hypothesized model, using recommendations from Hu & Bentler (1999). I found support for the hypothesized three-factor model, χ^2 (74) = 220.91, $p < .001$, CFI = .97, TLI = .96, RMSEA = .08, SRMR = .03. To examine *Hypothesis 1* (i.e., a three-factor model of the ILQ in which the items are set to load onto their respective factors ("facilitating belongingness" [A], "valuing uniqueness" [B], and "preventing exclusion" [C]), will fit significantly better than a one-factor model (*H1a*), a two-factor model with A and B together (*H1b*), a two-factor model with A and C together (*H1c*), and a two-factor model with B and C together (*H1d*)), I then tested a series of alternative models including two- and one-factor models. Based on a series of chi-square difference tests, I found that each of these models fit significantly worse than the hypothesized model, supporting *Hypothesis 1* (see model comparisons in Table 5).

To explore Research Question 1 (i.e., whether second-order factor loadings are sufficiently large enough (i.e., .40 or higher) to support the idea of a common second-order factor of inclusive leadership, I specified a second-order factor onto which the first order factors were set to load, then ran a second-order factor CFA and examined the factor loadings. All three factors demonstrated relatively large factor loadings onto the second order factor ($\lambda_{belongingness}$ = .93, $\lambda_{uniqueness}$ = .95, $\lambda_{mistreatment}$ = .86), indicating strong

support for a second-order factor of “inclusive leadership." I therefore used summated scores from the full scale in future analyses.

Phase 5: Convergent and Discriminant Validity

After obtaining validity evidence for the factor structure of my items, I proceeded to assess convergent validity (the extent to which my measure correlates with other theoretically similar measures) and discriminant validity (the extent to which my measure does not correlate with theoretically dissimilar measures). This phase is important for gathering additional evidence of construct validity (Hinkin, 1998).

Sample. I again utilized CloudResearch to recruit a sample of 1,443 working adults from Amazon MTurk and used the same screening technique as prior phases (Smith et al., 2015). I removed 635 participants who did not meet the inclusion criteria of being currently employed and working at least 20 hours per week under the direction of a supervisor and 536 who failed attention check items (i.e., “Please select Never for this one.”). I removed an additional 17 cases that were missing at least 50% of their data, resulting in a final sample of 255 participants (M_{Age} = 38.98; SD_{Age} = 11.52; 57% Women; 51% White; 28% Black/African American; 11% Hispanic/Latinx; 7% Asian; 95% full-time employed). This sample size is in line with recommendations by Schmidt and Stults (1986).

Design. Participants were presented with initial screening items as in prior phases (Smith et al., 2015). Those who met all criteria proceeded to the full survey, which included 10 scales, including the newly developed ILQ (see Appendix B for the complete

list of items). I provided participants with the same instructions as in the exploratory and confirmatory factor analysis stages. Participants also provided demographic information.

Convergent validity measures. To assess convergent validity, participants responded to two existing measures of inclusive leadership and measures of similar constructs including supervisor support, transformational leadership, abusive supervision, and leader allyship behaviors. Unless otherwise noted, I provided participants with the following instructions for each scale: "Please indicate how often your direct supervisor (manager) behaves in the following ways," and then asked them to respond to each item on a 5-point frequency scale (1 = 'Never', 5 = 'Frequently, if not always').

Inclusive leadership. I utilized the two most commonly used existing measures of inclusive leadership (Nembhard & Edmondson, 2006; Carmeli et al., 2010). Nembhard and Edmondson's (2006) scale includes three items measuring the degree to which leaders invite and appreciate others' contributions. A sample item includes "My supervisor asks for the input of all team members." The measure demonstrated strong reliability (α = 93). Carmeli et al.'s (2010) scale includes nine items measuring a leader's openness, accessibility, and availability. A sample item includes "My manager is open to hearing new ideas." The scale demonstrated strong reliability (α = .96).

Perceived supervisor support. I measured perceived supervisor support using an adapted version of the Survey of Perceived Organizational Support (SPOS; Eisenberger et al., 1986), replacing the word "organization" with "supervisor," in line with previous research (e.g., Eisenberger, 2002; Kottke & Sharafinski, 1988; Hutchison, 1997; Rhoades et al., 2001). Specifically, I adapted four items (Items 9, 10, 25, and 35; Eisenberger et

al., 1986) on the basis of their high factor loadings. I provided participants with the following instructions: “Please indicate your level of agreement with each statement about your immediate supervisor,” and then asked them to respond to each item on a 5-point agreement scale (1 = ‘Agree not at all’, 5 = ‘Strongly agree). A sample item includes “My supervisor cares about my opinions.” The scale demonstrated strong reliability (α = .95).

Transformational leadership. I used the 7-item Short Measure of Transformational Leadership (Carless et al., 2000). A sample item includes “My supervisor gives encouragement and recognition to staff.” The scale demonstrated strong reliability (α = .96).

Abusive supervision. I measured abusive supervision using six items from Tepper’s (2000) 15-item abusive supervision scale, in line with past research that tested the content validity of a shortened scale (Harris et al., 2011; α = .92). A sample item includes “My supervisor tells me my thoughts or feelings are stupid.” The scale demonstrated strong reliability (α = .91).

Allyship. I will assess allyship using the 7-item Ally Support and Advocacy Measure, which measures individual’s allyship behaviors along the dimensions of support and advocacy (Snoeyink et al., 2019). A sample item includes “My supervisor reminds others to use inclusive language.” I provided participants with the following instructions: “Please indicate your level of agreement with each statement about your immediate supervisor,” and then asked them to respond to each item on a 5-point

agreement scale (1 = 'Agree not at all', 5 = Strongly agree). The scale demonstrated good reliability ($\alpha_{support}$ = .89, $\alpha_{advocacy}$ = .86).

Discriminant validity measures. To assess discriminant validity, participants responded to measures of three constructs: including adventurousness, humor, and orderliness. For all three measures, I provided participants with the following prompt: "To what degree do you agree the following statements describe your supervisor's personality?" and then asked them to respond to each item on a 5-point agreement scale (1 = 'Agree not at all', 5 = 'Strongly agree). See Appendix B for the complete list of items.

Adventurousness. I assessed adventurousness using five items of a 10-item adventurousness scale from the revised NEO Personality Inventory (NEO-PI-R: Costa & McCrae, 1992). A sample item includes "My supervisor prefers variety to routine." This scale demonstrated good reliability (α = .89).

Humor. I assessed humor using five items from Peterson and Seligman (2004). A sample item includes "My supervisor goes out of their way to make people laugh." This scale demonstrated strong reliability (α = .96).

Orderliness. I assessed orderliness using five items of a 10-item orderliness scale from the revised NEO Personality Inventory (NEO-PI-R: Costa & McCrae, 1992). A sample item includes "My supervisor loves order and regularity." This scale demonstrated good reliability (α = .88).

Results. To test *Hypothesis 2*, (that the ILQ will demonstrate convergent validity (moderately correlate) with existing measures of inclusive leadership (*H2a*), perceived

supervisor support (*H2b*), transformational leadership *(H2c),* allyship (*H2d*), and abusive supervision (*H2e*)) and *Hypothesis 3* (that the ILQ will demonstrate discriminant validity (low correlations) with measures of adventurousness (*H3a*), humor (*H3b*), and orderliness (*H3c*), I obtained Pearson's *r* correlations by correlating the newly developed scale with the other measures and then examining the magnitude and significance of correlations.

In support of *Hypothesis 2*, the ILQ dimensions and full scale scores significantly correlated with all convergent validity measures, including both Carmeli et al.'s (2010) measure (r = .76 – .83) and Nembhard and Edmondson's (2006) measure (r = .78 – .86) of inclusive leadership (*H2a),* as well as supervisor support (r = .74 – .80, *H2b),* transformational leadership (r = .75 – .82, *H2c*), allyship (r = .64 – .82, *H2d)*, and abusive supervision (r = -.50 – -.54, *H2e).*

Hypothesis 3 was not supported. I found that the ILQ and its subscales significantly correlated with all discriminant validity variables to a moderate or large extent, including adventurousness (r = .52 –.60, *H3a),* humor (r = .66 –.73, *H3b),* and orderliness (r = .32 – .42, *H3c).*

Phase 6: Predictive Validity/Incremental Validity

To gain additional construct validity evidence, I then examined criterion related validity of the ILQ for a measure of inclusion climate perceptions. I also assessed the incremental validity of the ILQ in predicting inclusion climate over and above existing measures of inclusive leadership (Carmeli et al., 2010; Nembhard & Edmondson, 2006). Theoretically, the ILQ should incrementally predict inclusion climate perceptions

because it is based on a broader understanding of inclusive leader behaviors (i.e., preventing mistreatment and discrimination in addition to facilitating belonging and supporting employees' unique identities) than existing measures.

Sample. I recruited a total of 1,194 working adults from various sources. Specifically, I utilized convenience and snowball sampling to collect a total of 1,095 responses and recruited 99 working adult students from a university in the Northwest United States. As an incentive, participants were provided a link at the end of the survey to enter to win one of five $100 Amazon gift cards. The student sample was also offered extra credit in exchange for participating in the study. I again used Smith et al.'s (2015) screening technique as well as instructed response attention check items (i.e., "Please select "Never" for this one"). After removing cases that failed at least one attention check item, did not meet inclusion criteria, and/or were not based in English-speaking countries, a final sample of 482 participants remained (n=482). Previous incremental validity studies have demonstrated significant change in R^2 values of .02. Based on that, a power analysis indicated I would need 352 participants to detect a small effect at 80% power. I therefore utilized 352 randomly selected cases from the data set (M_{Age} = 35.4, SD_{Age} = 9.75; 63% Women; 83% White; 86% employed full-time), preserving the remaining cases (n = 130) for a separate study.

Design. After responding to screening questions, participants were instructed to rate their immediate supervisor using the newly developed ILQ and the two most commonly used existing measures of inclusive leadership (Carmeli et al., 2010; Nembhard & Edmondson, 2006). They were also asked to respond to items measuring

their inclusion climate perceptions, job satisfaction, affective commitment, turnover intentions, and experiences with incivility, followed by several demographic items.

Measures. In order to assess the predictive and incremental validity of the ILQ, participants provided responses to the ILQ, two existing measures of inclusive leadership, and five job outcomes (i.e., inclusion climate perceptions, job satisfaction, affective commitment, turnover intentions, and incivility). Unless otherwise indicated, participants responded to all items using a 5-point agreement scale (1 = 'Agree not at all', 5 = 'Strongly agree'). See Appendix B for a full list of items for each scale.

Inclusion Climate Perceptions. I measured inclusion climate perceptions using six items from the 10-item Work Group Inclusion scale (Chung et al., 2020). I used the three highest-loading items of each content area of the scale (i.e., belongingness and uniqueness). Participants were instructed to indicate the degree to which they agree with each of the following statements about their work group (i.e., immediate team). A sample item includes "I am treated as a valued member of my work group." This scale demonstrated strong reliability (α = .87).

Job Satisfaction. I assessed job satisfaction using the three-item Michigan Organizational Assessment Questionnaire Job Satisfaction Subscale (Cammann, Fichman, Jenkins, & Klesh, 1979; 1983). I adjusted one item to remove reverse scoring (i.e., "In general, I don't like my job" to "In general, I like my job."). This scale demonstrated good reliability (α = .89).

Affective Commitment. I measured affective commitment using three items from Allen and Meyer's (1990) 8-item Affective Commitment subscale, which assesses

employees' emotional attachment to and identification with the organization. I used the three highest-loading items of the scale. A sample item includes, "This organization has a great deal of personal meaning for me." The scale demonstrated strong reliability (α = .90).

Incivility. I measured coworker incivility using four items from Cortina and colleagues' (2001) 7-item Workplace Incivility Scale (WIS; Cortina et al., 2001), which asks participants to indicate how frequently they have experienced a list of uncivil behaviors from a coworker in the last month. I used the four items that aligned most with mistreatment. A sample item includes, "Over the last month, a coworker made derogatory remarks about me." The scale demonstrated good reliability (α = .86).

Turnover Intentions. I assessed turnover intentions using one item: "I will probably look for a new job during the next year" (Porter et al., 1976).

Inclusive Leadership. I utilized the same two measures of inclusive leadership to assess incremental validity that I used in my convergent validity stage. Both Carmeli et al. (2010; α = .94) and Nembhard and Edmonson (2006; (α = .91) demonstrated good reliability.

Results. I conducted all statistical analyses in R. I first assessed whether assumptions have been met. First, there were only outliers on demographic variables (i.e., organizational tenure and tenure with one's manager). I conducted a sensitivity analysis and found no differences in scale reliabilities, correlations or regression results, so retained those outliers for all analyses. Second, scatter plots revealed linear associations between the ILQ and all outcome variables. ILQ scores were significantly and positively

related to inclusion climate (r = .64, p < .001), job satisfaction (r = .56, p < .001), and affective commitment (r = .46, p < .001), and negatively related to incivility (r = -.34, p < .001) and turnover intentions (r = -.34, p < .001). Third, all focal variable residuals were normally distributed. See a correlation table for all variables in Table 7.

To test Hypothesis 4 (that the ILQ will predict inclusion climate perceptions (*H4a*), job satisfaction (*H4b*), affective commitment (*H4c*), incivility (*H4d*) and turnover intentions (*H4e*), I conducted a series of linear regression analyses. *H4a-H4e* were all supported. Inclusive leadership significantly positively predicted inclusion climate perceptions (b = .60, p<.001, CI[.53, .68]), job satisfaction ratings (b = .64, p<.001, CI[.54, .74]), and affective commitment (b = .60, p<.001, CI[.48, .72]), and significantly negatively predicted incivility (b = - .36, p<.001, CI[-.44, -.24]) and turnover intentions (b = -.52, p<.001, CI[-.69, -.35]). See Figure 2.

To test Hypothesis 5 (that the ILQ would have significant incremental validity in the prediction of inclusion climate above and beyond two existing measures of inclusive leadership), I conducted a series of two separate hierarchical regression models with inclusion climate perceptions as the dependent variable (i.e., one for each of the existing measures of inclusive leadership; Carmeli et al., 2010 (*H5a*); Nembhard & Edmondson, 1996 (*H5b*)). For each model,

I first entered scores from an existing inclusive leadership measure as a predictor of the outcome, then added ILQ scores as a second predictor, fit a new model, and compared the models by examining changes in R^2. I ensured no multicollinearity was present among predictors (i.e., all tolerance statistics [1-R^2] were above .20 and VIF values were below

5). To support *H5*, I expected the ΔR^2 to be significant for the ILQ. *H5a* and *H5b* were both supported (see Table 8).

First, Carmeli et al.'s (2010) measure and the ILQ explained 47% of the variance in inclusion climate perceptions, $F(2, 349)=151.5$, $p<.001$. Carmeli et al.'s measure was significant at step 1 ($b= .36$, $p<.01$), and in support of *H5a*, the ILQ was significantly related to inclusion climate over and above that measure ($\Delta R^2 =.05$, $p < .01$). Additionally, Nembhard and Edmondson's (2006) scale and the ILQ explained 44% of the variability in inclusion climate, $F(2,349) =137.8$, $p<.001$. Nembhard and Edmondson's (2006) measure was also significant at step 1 ($b=.37$, $p<.01$), and in support of *H5b*, the ILQ was significantly related to inclusion climate over and above that measure ($\Delta R^2 =.07$, $p < .01$). Therefore, although the ILQ is conceptually related to these two existing scales, it explains unique variance in the outcome of inclusion climate.

Exploratory analyses results. To analyze Research Question 2 regarding whether the ILQ would more strongly predict job satisfaction (*RQ2a*), commitment (*RQ2b*), incivility (*RQ2c*) and turnover intentions (*RQ2d*) compared to existing measures of inclusive leadership, I ran separate regression analyses in the same manner as I did to test Hypothesis 4. Specifically, I ran eight separate hierarchical regression models (i.e., one for each of the existing measures of inclusive leadership on each of the four outcomes). The ILQ demonstrated incremental validity beyond both Carmeli et al.'s (2010) and Nembhard and Edmondson's (2006) scales for job satisfaction (ΔR^2s=.05 and .06 respectively), affective commitment (ΔR^2s=.08 and .07 respectively), and turnover intentions (ΔR^2s=.01 and .02 respectively). The ILQ did not demonstrate incremental

validity for incivility beyond either Carmeli et al. or Nembhard and Edmondson's scales. See Table 8 for detailed results.

Finally, to analyze Research Question 3 regarding whether individuals from traditionally marginalized groups (i.e., on the basis of race and gender) would interpret and react differently to inclusive leadership behaviors (i.e., respond differently to perceptions of inclusion climate) than those from majority groups, I ran separate moderated multiple linear regression (MMLR) models in R.

First, because there were higher proportions of White individuals (83%) and women (63%) in my sample, I created two new versions of the dataset: one with equal numbers male (n_{male}= 120) and "non-male" (n =120; n_{women}= 115, $n_{transgender}$ =3, $n_{non-binary}$=2) participants, and one with equal numbers of people who only identified as White (n= 82) and people who identified as one or more of the race/ethnicity categories but did not also identify as White (n =82; n_{Black}= 23, n_{Latinx} =26, n_{Asian}= 14, n_{Native}= 9; $n_{Self\ Describe}$= 10; $n_{MiddleEastern}$= 4; $n_{PacificIslander}$= 1). To do this, I created two new variables, Race and Gender, and coded all White and Male participants as 0 and all other cases as 1 for those variables. I then counted how many "non-White" participants there were in the full data set and used a random number generator to randomly select the same number of White participants, and saved those cases into a new data set. I followed the same steps to select equal numbers of Male and "Non-Male" participants.

I then utilized moderated multiple linear regression to analyze whether both gender and race moderated the impact of inclusive leader behaviors on inclusion climate perceptions such that the predictions significantly differed for members of majority

groups (i.e., men, Whites) from those of gender and racial minority groups. In order to improve the interpretability of the main effects, I centered the dependent variable of inclusion climate perceptions.

To test whether Race moderated the impact of inclusive leadership on inclusion climate perceptions, I first regressed inclusion climate perceptions onto the ILQ and found ILQ scores were once again positively and significantly associated with inclusion climate perceptions (b = .59, p < .001, CI[.47,.71]). In the second step, I added Race to the model as a moderator to investigate whether intercept differences exist, and found no significant differences (b=-.10, *ns*). In the third step, I added the interaction term between the ILQ and Race, resulting in a moderated multiple linear regression model, and found that race did not moderate the association between ILQ scores and inclusion climate perception scores to a statistically significant extent (b = -.18, *ns*). To explore further, I repeated these steps for each sub-dimension of the ILQ and found that race did not moderate the association between inclusion climate perception scores and any of the ILQ sub-dimensions of facilitating belongingness (b = -.12, *ns*), supporting uniqueness (b = -.13, *ns*), or preventing mistreatment ratings (b = -.18, *ns*).

To test whether Gender (Male vs. Non-Male) moderated the impact of inclusive leadership on inclusion climate perceptions, I once again started by regressing inclusion climate perceptions onto the ILQ. In the second step, I added Gender to the model as a moderator to investigate whether intercept differences exist, and found no significant differences, b = -.09, *ns*. In the third step, I added the interaction term between the ILQ and Gender, resulting in a moderated multiple linear regression model, and found that

gender did not moderate the association between ILQ scores and inclusion climate perception scores to a statistically significant extent (*b* = .12, *ns*). To explore this further, I repeated these steps for each sub-dimension of the ILQ and found that gender did not moderate the association between inclusion climate perception scores and any of the ILQ subdimensions of facilitating belongingness (*b* = .58, *ns*), supporting uniqueness (*b* = .79, *ns*), or preventing mistreatment ratings (*b* = .14, *ns*).

Chapter 4: Discussion

Overview

Many organizations view diversity as a strategic business priority but struggle to reach their goals related to the retention, promotion, or full participation of their employees, particularly those from under-represented or traditionally marginalized groups. Ample empirical evidence suggests that employers may enhance outcomes related to work team diversity by fostering positive inclusion climate perceptions (Chrobot-Mason & Aramovich, 2013; Jackson & Joshi, 2011; Jehn & Bezrukova, 2010). Managers are critical to creating inclusive climates because they represent the organization, set and enforce policy, and shape culture. To better understand which behaviors shape employee perceptions of inclusion climate and lead to other positive job outcomes, we must have clarity on the construct of inclusive leadership itself. To date, however, scholars have operationalized and measured the construct in multiple ways that do not capture the full content domain of inclusion as defined in the diversity and inclusion literature. With a theoretically based and empirically validated measurement tool, it is possible to better understand the impact of managers' inclusive behaviors, which will inform future efforts to enhance inclusion for all employees and reduce discrimination of individuals from stigmatized groups.

Across seven samples and 1,224 participants, I developed and examined the utility of a multidimensional conceptualization of inclusive leadership. To do this, I relied upon established theory on workplace inclusion (Shore et al., 2011) and integrated inclusive leadership theory from Randel et al. (2018) and Perry et al. (2020) in order to

develop the three dimension ILQ. Additionally, I assessed the new measure's predictive and incremental validity above and beyond existing scales of inclusive leadership, which do not incorporate important aspects of effective diversity management. Finally, I explored whether inclusive leadership has a differential impact on people from traditionally more privileged groups compared to traditionally marginalized groups (i.e., along the dimensions of gender and race).

Findings

Using a deductive scale development approach (Hinkin, 1998), I developed the ILQ based upon recent inclusive leadership theory by Randel et al. (2018), who proposed inclusive leadership was comprised of the two dimensions of facilitating belongingness and supporting uniqueness, and Perry et al. (2020), who proposed inclusive leaders must also work to address and prevent mistreatment. Specifically, SMEs developed a pool of items that were evaluated for content validity, trimmed further in a series of exploratory factor analyses, and then assessed for factor structure using confirmatory factor analysis. In support of *Hypothesis 1*, I found that the three-factor model fit better than any examined one-factor (*H1a*) or two-factor models (*H1b*). This is theoretically important because these findings provide strong support for a theoretically grounded conceptualization of inclusive leadership based on the predominant model of inclusion in the diversity and inclusion literature (i.e., Randel et al., 2018; Shore et al., 2011). By doing so, this model redefines the inclusive leadership construct, which has been misconceptualized within the diversity and inclusion literature. Past conceptualizations of inclusive leadership (i.e., Carmeli et al., 2010, Nembhard & Edmondson, 2006) narrowly

define inclusive leadership as being "open" and "accessible" in order to facilitate psychological safety and creativity for all employees, and that limits our understanding of what behaviors contribute to employees' perceptions of being included at work. This research realigns the construct with our current understanding of what it means for organizations and individuals to be "inclusive" (i.e., facilitating employees' belonging and supporting their uniqueness while also preventing mistreatment and exclusion). After establishing the ILQ's factor structure, I confirmed that a second-order factor of inclusive leadership exists. This is significant from a theoretical perspective because it demonstrates support for a latent construct of inclusive leadership. It is also significant from a practical perspective because it means overall ILQ scores can be used in research and practice.

I then assessed the convergent (*Hypothesis 2*) and discriminant (*Hypothesis 3*) validity of the ILQ with multiple similar and disparate constructs using summated ILQ scores from a new sample. The ILQ demonstrated convergent validity as expected, positively correlating with two existing measures of inclusive leadership, supervisor support, transformational leadership, and allyship, and negatively correlating with abusive supervision. All correlations were large (*rs* = |.54| - |.86|), with the highest correlations among ILQ and other measures of leadership styles: Nembhard and Edmondson's inclusive leadership scale (*r* =.86), Carmeli et al.'s inclusive leadership scale (*r* =.83), and transformational leadership (*r* =.82). I also examined these correlations at the ILQ factor level and found correlations remained strong with all variables (*rs* = |.50| - |.83|).

To examine *Hypothesis 3* and gather evidence of discriminant validity, I examined the correlations between the ILQ and measures of humor, adventurousness, and orderliness Counter to what I expected, I found that all three variables correlated with the ILQ (*H3a-H3c* unsupported). Humor correlated the strongest with the ILQ and its sub dimensions (*rs* = .66-.73), followed by adventurousness (*rs* = .52-.60), and orderliness (*rs* = .32-.42). One possible explanation for the strong correlations between the ILQ and all three variables is common method bias. Because all variables were included in the same survey instrument, common method bias could have contributed to inflated correlation coefficients, thereby reducing the discriminant validity of the scale (Podsakoff et al., 2012).

These strong correlations could also be explained by a halo effect. All three discriminant validity variables I utilized are considered positive characteristics in American society. People who have overall positive feelings toward their supervisor might rate their supervisor high on humor, adventurousness, and orderliness because they like them and believe they would have these positive qualities.

It could also be the case that people who are perceived as being humorous are also people who others perceive to be welcoming and inclusive, which is a key part of inclusive leadership. Similarly, people who are seen as adventurousness may also be perceived as open-minded, which might relate to perceptions of overall inclusivity. Finally, perhaps perceptions of people as orderly would relate to perceptions of them as being fair-minded and equitable.

After assessing the convergent and discriminant validity of the ILQ, I examined its predictive validity for inclusion climate perceptions (*H4a*), job satisfaction (*H4b*), affective commitment (*H4c*), incivility (*H4d*), and turnover intentions (*H4e*). I found that the ILQ (using summated ILQ scores), significantly predicted all outcomes; *Hypothesis 4a-e* supported). These findings are not surprising because research indicates that when employees feel their organizations have effectively managed diversity and fostered a climate of inclusion, they report higher levels of job satisfaction and affective commitment (Acquavita et al., 2009; Mor Barak et al., 2016). Further, past research indicates that inclusive leadership negatively relates to turnover intentions (Nishii & Mayer, 2009). Based on that and other research, it is likely that inclusion climate mediates the impact of inclusive leadership on job satisfaction, commitment and turnover, though such analyses were outside the scope of this dissertation. Additionally, research has found that positive leader behaviors can lead to lower levels of incivility by enhancing employees' perceptions of norms for respect (Walsh et al., 2017).

Regarding incremental validity, the ILQ explained additional variance in inclusion climate above and beyond existing measures of inclusive leadership as predicted (*H5* supported). Specifically, the ILQ explained an additional five percent of variance (R^2=.05, p<.01) over and above Carmeli et al.'s measure and an additional seven percent of variance (R^2=.07, p<.01) over and above Nembhard and Edmondson's measure. This is not surprising given that two of the ILQ's dimensions emphasize facilitating belonging and supporting uniqueness (based on Randel et al.'s theoretical framework), which directly map onto the conceptualization of inclusion climate

perceptions as involving a sense of belonging and feeling that one's uniqueness is valued (Chung et al., 2020; Shore et al., 2011). Additionally, correlations among ILQ subdimensions and the existing measures indicate that Carmeli et al.'s and Nembhard and Edmondson's measures correlate relatively lower with the ILQ's "preventing mistreatment" dimension (*rs* =.66 and .67 respectively), compared to "facilitating belongingness" (*rs* =.81 and .77) and "supporting uniqueness" (*rs* =.72 and .70). This indicates that adding a third dimension to Randel et al.'s (2018) framework that emphasizes the prevention of mistreatment and exclusion helps explain the impact of leadership behaviors on inclusion climate. Overall, these findings bolster a three dimensional conceptualization of inclusive leadership that integrates Randel et al.'s (2018) and Perry et al.'s (2020) frameworks.

Additionally, I explored whether the ILQ would explain additional variance in other job outcomes above above and beyond existing measures of inclusive leadership: job satisfaction (*RQ2a*), commitment (*RQ2b)*, incivility (*RQ2c)*, and turnover intentions (*RQ2d).* Overall, the ILQ explained a significant amount of additional variance over and above all outcomes except incivility.

In terms of incivility, although the ILQ significantly predicted incivility (r = -.34, $p < .001$), it did not explain additional variance over and above existing measures. Both Carmeli et al.'s and Nembhard and Edmondson's measures already demonstrated a moderately strong relationship with incivility (R^2s= .17 and .14 respectively), which is somewhat surprising given neither measure includes items that cover leader behaviors involving promoting positive relationships among coworkers or preventing mistreatment

and incivility was measured as how frequently respondents had been the target of uncivil behaviors over the last month (e.g, "Over the last month, a coworker made demeaning or derogatory remarks about me"). Given that the dependent variable measures were placed directly following the existing measures in the survey, perhaps common method bias inflated the correlation coefficients, leading to erroneous results. Another explanation could be that I instructed participants to indicate the frequency by which they had experienced certain behaviors in the last month. This likely curtailed variability that I would have captured had I not specified a timeframe.

Finally, I explored whether inclusive leadership would have a different level of impact on inclusion climate perceptions for people from various groups (i.e., on the basis of race and gender). I didn't hypothesize a specific direction because some past research suggests under-represented or traditionally marginalized group members may have stronger reactions to inclusive leaders, while other research suggests majority group members might (e.g., backlash). Nonetheless, I expected the effect to be stronger for the under-represented or marginalized groups because literature has shown that, on average, people from under-represented or traditionally marginalized group members rate their workplaces as less inclusive (Kossek & Zonia, 1993; Mor Barak et al., 1998). However, I did not find significant slope differences based on race or gender. These findings indicate that when leaders are perceived as acting in an inclusive manner, their behaviors will impact inclusion climate perceptions for all team members by the same degree. This is theoretically and practically meaningful because it suggests that a certain set of manager behaviors can drive a positive inclusion climate for all employees. The fact that there

were not significant slope differences between the two sets of groups runs somewhat counter to research showing that people from majority groups tend to rate their workplaces as more inclusive.

Theoretical Implications

The results of this study have several theoretical implications. A key contribution of this dissertation is the theoretical refinement of the inclusive leadership construct. As previously discussed, earlier research on inclusive leadership is relatively young and has conceptualized the construct inconsistently. Scholars have operationalized it as inviting others' contributions (Nembhard & Edmondson, 2006); as leader openness, availability, and accessibility (Carmeli et al., 2010); or as group-level LMX (Nishii et al., 2009). More recent theory has defined the construct as behaviors that fulfill employees' needs for belongingness and uniqueness (Randel et al., 2018), aligning with the predominant definition of workplace inclusion (i.e., Shore et al., 2011), which is based on ODT (Brewer, 1991). Other recent theory and research (Perry et al., 2020) suggests inclusive leadership also involves preventing exclusion, particularly of employees with marginalized identities. Across multiple studies, I provided empirical evidence for a multi-dimensional conceptualization of inclusive leadership based on this current theoretical understanding and demonstrated that inclusive leadership contributes to important employee outcomes.

A second contribution of this research is the development and empirical validation of a measure. In doing so, this work also extends prior research and addresses calls in the organizational literature for such a tool (e.g., Chung et al., 2020; Randel et al.,

2018; Shore et al., 2018; Zhang et al., 2016) and provides a theoretically grounded foundation for the burgeoning literature on inclusive leadership.

Finally, this research contributes to and integrates the diversity and leadership literatures by developing and validating a set of leadership behaviors that foster a positive inclusion climate, which should drive other positive outcomes for employees including job satisfaction, commitment, reduced mistreatment and turnover, and a host of other outcomes. This work also helps to establish inclusive leadership as a distinct style of leadership.

Practical Implications

This study also has a number of practical implications. HR and business leaders are highly concerned with attracting and retaining diverse talent (Sigelman & Taylor, 2021), and to do this they must foster inclusive workplaces. This involves understanding what behaviors managers can perform to foster a greater sense of inclusion on their teams. The ILQ can help organizations achieve these goals. For example, the ILQ could be used in surveys to gain a sense of employee perceptions of their managers' inclusiveness. Results of such surveys could help drive targeted developmental interventions for managers. Being able to provide managers with feedback on the most effective behaviors they could perform to foster a positive inclusion climate could provide many benefits to organizations, from increased engagement, commitment, performance, and well-being to reduced turnover. The ILQ could also be used to develop an inclusiveness competency for managers to be used as part of an individual development plan.

Second, the ILQ could also be used to design and evaluate the effectiveness of an inclusive leadership training program for managers. Inclusive leadership models are proliferating in consulting firms, suggesting large enterprises are seeking information on how to develop these skills among their managers. The ILQ dimensions and items could be used to develop training content and the ILQ could be used to evaluate managerial effectiveness before and after the training. Without proper measurement, organizations might evaluate programs ineffectively.

Limitations and Future Research

As with any study, this work has a number of potential limitations. First, I collected data for the exploratory factor analysis, confirmatory factor analysis, and convergent and discriminant validity stages from one pool of participants, workers on Amazon's MTurk. Many consider this a limitation due to data quality concerns (i.e., bots). Recent research, however, has demonstrated the utility and validity of MTurk workers as a participant pool (Buhrmeister et al., 2011) and screening protocols following best practices were utilized to ensure quality data from that source (Smith et al., 2015). Responses from MTurk workers have demonstrated similar data quality as workplace samples when collected in the U.S. (Feitosa et al., 2015). Additionally, my use of Cloud Research to tailor my recruiting helped ensure higher quality data.

Second, my predictive/incremental validity phase included both the predictor variables and outcome measures via a cross sectional design, rendering directionality unclear. However, inclusion climate scores should theoretically be predicted by leader behaviors, and past research has demonstrated that leader behaviors are a primary driver

of organizational climate (Schneider et al., 2017). This design, then, may serve as a beneficial first step for assessing the predictive and incremental validity of the ILQ.

Third, for my exploration of Research Question 2 (how identity characteristics such as race and gender impact the interpretation and reactions to inclusive leadership behaviors), I compared only Whites vs. non-Whites and men vs. all individuals who identify as women or another gender, a similar approach as in past research (e.g., King et al., 2012). However, this method ignores differences that may exist among different groups. Although no significant differences were found between Whites and Non-Whites, future research should consider differences between minority groups.

There are several avenues for future research. First, future research should continue to examine the ILQ's validity with additional research designs. For example, my predictive study findings should be triangulated with a longitudinal study so as to better test the directionality of my findings and assess potential mediating mechanisms (e.g., inclusive leadership could predict inclusion climate, which could subsequently predict job outcomes). Additionally, qualitative research that confirms employees expect inclusive leaders to behave in ways identified by my measure, would ground my measure in real-life experiences. It is particularly important to include qualitative research from those from under-represented and marginalized groups, as more work is needed in understanding the experiences of women and people of color with inclusive leadership (Bilimoria et al., 2008; Shore et al., 2018). Second, future research should focus on the development and evaluation of organizational interventions and training programs based on the ILQ's dimensions of facilitating belongingness, valuing uniqueness, and

preventing exclusion. Using an intervention design, where inclusive leadership is a pre- and post-measure to a leadership training initiative focused on developing these behaviors in managers would be a useful additional way to assess the utility of the measure and its reliability. Past research has demonstrated that managers can be effectively trained to provide more support for their employees by engaging in specific behaviors, which in turn can enhance employee performance and well-being (Barling et al., 1996; Dimoff & Kelloway, 2019; Hammer et al., 2016; Hammer et al., 2019; Kelloway et al., 2000). Third, future research could also directly compare inclusive leadership to other forms of leadership to examine its comparative utility in predicting inclusion climate and other outcomes. Fourth, validating a short version of the measure (e.g., three items) would enhance the practicality of the measure for use in organizational surveys and research studies. Fifth, future studies could explore whether, in addition to race and gender, employees of additional groups respond to the scale differently (i.e., ability, sexual orientation, gender identity). Sixth, future research could examine the relative importance of the sub-dimensions of inclusive leadership behaviors on inclusion climate perceptions for employees of various groups. It could be the case that preventing mistreatment is a relatively stronger driver than fostering feelings of belonging or uniqueness. Seventh and finally, future research should examine to what degree personal characteristics of the leader play a role in follower perceptions. For example, it is reasonable to assume that people who hold marginalized identities (i.e., women, racial minorities, and members of other stigmatized groups) would be rated higher on the ILQ by others because they have insider knowledge on what behaviors should enhance

follower perceptions of inclusion. These effects would likely be dependent on other contextual factors, however; for example, research has found that leader-follower gender dissimilarity is negatively associated with perceptions of inclusion and the negative relationship is stronger for men than for women (Bae et al., 2017). More research is needed to better understand these relationships.

Conclusion

This scale development and validation research represents an effort to clarify the construct of inclusive leadership and confirm its impact on relevant workplace outcomes. Although many organizations view diversity as a strategic business and HR priority, many struggle to effectively manage diverse teams and retain talent due to a lack of inclusion. By providing clarity about the inclusive leadership construct, developing and validating a tool assessing inclusive leader behaviors, and testing how employee ratings of their leaders are related to employee and organizational outcomes, the field can create better interventions for managers to develop these skills that will help them enhance inclusion climate perceptions within their organizations.

Importantly, this research provides empirical evidence for a multi-dimensional conceptualization of inclusive leadership that supports Randel et al.'s (2018) proposed framework and adds a third dimension ("preventing mistreatment"; Perry et al., 2020). The strong support for this third dimension demonstrates that leaders need to simultaneously promote inclusion and prevent exclusion to foster inclusive work environments and that each may require different behaviors. I hope this research spurs further examination of this important phenomenon.

Tables

Table 1

Study Design

Phase	Purpose
Phase 1: Item Generation	Develop initial pool of items based on Randel et al.'s (2018) and Perry et al.'s (2020) inclusive leadership dimensions
Phase 2: Content Validation / Item Reduction	
Content Adequacy	Determine items' appropriateness with a sorting task in which participants identified which of Randel et al.'s (2018) and Perry et al.'s (2020) dimensions best represented each item
Item Reduction	Evaluate clarity and relevance of remaining items
Phase 3: Exploratory Factor Analysis	Refine and reduce number of items, ensuring strong loadings on their hypothesized factor and minimal cross-loadings
Phase 4: Confirmatory Factor Analysis	Confirm factor structure of final item pool
Phase 5: Convergent / Discriminant Validity	Establish nomological network of developed factors with theoretically similar and distinct constructs
Phase 6: Predictive / Incremental Validity	Evaluate predictive and incremental ability of developed factors

Table 2

Inclusive leadership compared to other leadership styles

Leader Style	Differentiation
Transformational leadership	Focuses on motivating and developing employees to meet organizational objectives. Inclusive leadership emphasizes accepting employees for who they are and helping them feel like they belong and can contribute their unique perspectives and abilities.
Authentic leadership	Focuses on leader actions and behaviors that reflect their authentic beliefs and self in their interactions with others and, as such, is more leader-focused than inclusive leadership.
Leader-member exchange (LMX)	Inclusive leadership goes beyond individual relationships to create an inclusive team environment where all team members feel they belong and their unique perspectives are valued.
Charismatic leadership	More leader-focused, whereas inclusive leaders start from supervisees' individual needs to belong and bring their full selves to work.
Ethical leadership	Emphasizes leaders' ethical behavior, which aims to inspire ethical behavior in followers; inclusive leaders focus on establishing respect within work teams by encouraging all team members to be themselves and preventing mistreatment.
Servant leadership	Focused on creating success for their followers, the organization, and other stakeholders such as customers and the broader community; inclusive leaders pay more attention to differences in employee needs.
Participative leadership	Emphasizes leader behaviors that involve followers in decision making, such as soliciting their suggestions and taking them into account; Inclusive leaders go beyond that by ensuring all team members' voices are heard, that they are treated fairly, and that they feel connected to others.

Table 3

Existing inclusive leadership measures

Authors	Deficiencies
Nembhard & Edmondson (2006)	● No validity evidence ● Highly contextualized to medical industry
Carmeli et al. (2010)	● Insufficient validity evidence (i.e., lack of 2nd sample and incremental validity)
Jin et al. (2017)	● No validity evidence
Zheng et al. (2017)	● No validity evidence
Panicker et al. (2018)	● No validity evidence
Fang et al. (2019)	● Insufficient validity evidence (i.e., lack of convergent/discriminant and incremental validity)
Ratcliff et al. (2018)	● Insufficient validity evidence (i.e., content validity only) ● Highly contextualized to military context

Table 4

Summary of Exploratory Factor Analysis Results

Item Wording	Facilitating Belongness	Supporting Uniqueness	Preventing/ Addressing Mistreatment
1. Makes new team members feel included.	**.83**		
2. Makes everyone feel like they belong.	**1.00**		
3. Ensures that everyone feels welcome.	**1.00**		
4. Asks for different points of view, especially from those who may typically be under-represented.	.19	**.44**	.22
5. Encourages people to be themselves at work.		**.87**	
6. Allows team members to express their identities however they feel comfortable.	-.14	**1.02**	
7. Appreciates our team's diversity.	.24	**.48**	.22
8. Speaks up when someone makes a disrespectful comment regarding marginalized groups.	-.19		**1.01**
9. Confronts discrimination when they see it.	-.11		**.97**
10. Stands up for employees who are mistreated.	.10		**.84**
11. Corrects stereotypical comments people make at work.		-.12	**.92**
12. Creates and/or enforces policies and procedures that reduce bias.	.11	.12	**.69**
13. Takes appropriate actions to protect the marginalized people in our team.	.10		**.76**
14. Utilizes feedback from marginalized employees in workplace decisions, policies, and procedures.	.14		**.76**

Note. Factor loadings over .40 appear in boldface. Factor loadings below .10 have been omitted.

Table 5

Confirmatory Factor Analysis Model Comparisons

Model			$\chi 2$	df	$\Delta\chi 2$	CFI	TLI	RMSEA	SRMR
3 Factors	A1	1,2,3	220.91	74		.97	.96	.08	.03
2 Factors	B1	1+2, 3	362.82	76	141.91**	.93	.92	.12	.04
	B2	1+3, 2	630.29	76	409.38**	.87	.84	.16	.06
	B3	1, 2+3	489.24	76	268.33**	.90	.88	.14	.06
1 Factor	C1	1+2+3	772.13	77	551.22**	.84	.81	.18	.06

Note: ** $p < .001$

Table 6.

Summary of Convergent/Discriminant Validity Results

	M	*SD*	1	2	3	4	5	6	7	8	9	10
1. Inclusive Leadership Questionnaire	3.63	1.07	(-.96)									
2. Leader Inclusivity (Nembhard & Edmondson)	3.73	1.16	.86**	(-.93)								
3. Inclusive Leadership (Carmeli et al.)	3.76	1.06	.83**	.85**	(.96)							
4. Supervisor Support	3.65	1.24	.80**	.81**	.83**	(.95)						
5. Transformational Leadership	3.57	1.14	.82**	.81**	.86**	.86**	(.96)					
6. Allyship	3.21	1.13	.82**	.75**	.77**	.80**	.82**	(.92)				
7. Abusive Supervision	1.40	0.67	-.54**	-.52**	-.53**	-.58**	-.56**	-.49**	(.91)			
8. Adventurousness	3.07	1.02	.60**	.60**	.60**	.60**	.67**	.67**	-.29**	(.89)		
9. Humor	3.08	1.31	.73**	.70**	.71**	.74**	.80**	.74**	-.47**	.73**	(.96)	
10. Orderliness	3.50	1.00	.40**	.42**	.45**	.37**	.44**	.44**	-.23**	.40**	.34**	(.88)

Note. M = mean, SD = standard deviation. Alpha reliabilities presented in parentheses on the diagonal.

** $p < .001$

Table 7

Means, standard deviations, and correlations with confidence intervals for Predictive/Incremental Validity study

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11
1. ILQ	3.70	0.88	(.95)										
2. $ILQ_{belonging}$	3.94	0.93	.86**	(.88)									
3. $ILQ_{uniqueness}$	3.74	0.91	.89**	.77**	(.83)								
4. $ILQ_{preventmistreat}$	3.57	1.01	.95**	.70**	.73**	(.93)							
5. Inclusion	3.73	0.84	.63**	.62**	.58**	.57**	(.87)						
6. Job Satisfaction	3.66	1.00	.56**	.56**	.50**	.50**	.66**	(.89)					
7. Commitment	3.22	1.15	.46**	.39**	.39**	.45**	.54**	.67**	(.90)				
8. Incivility	1.86	0.89	-.34**	-.42**	-.36**	-.24**	-.29**	-.24**	-.10	(.86)			
9. Turnover Int.	2.64	1.52	-.30**	-.36**	-.27**	-.25**	-.23**	-.44**	-.32**	.41**			
10. Nembhard et al.	3.73	0.94	.77**	.77**	.70**	.67**	.61**	.52**	.37**	-.38**	-.29**	(.91)	
11. Carmeli et al.	3.84	0.91	.77**	.81**	.72**	.66**	.65**	.55**	.36**	-.42**	-.30**	.87**	(.94)

Note. *M* and *SD* are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. The confidence interval is a plausible range of population correlations that could have caused the sample correlation (Cumming, 2014). * indicates $p < .05$. ** indicates $p < .01$.

Table 8

Summary of Incremental Validity Results

	Inclusion Climate		Job Satisfaction		Commitment		Incivility		Turnover	
Predictor	Step 1	Step 2	Step 1	Step 2	Step 1	Step 2	Step 1	Step 2	Step 1	Step 2
Carmeli et al., 2010	.60***	.36**	.61***	.32**	.46***	.02	-.41***	-.38**	-.49***	-.26*
ILQ		.32**		.38**		.58**		-.04		-.31*
R^2	.42***	.47***	.30***	.35***	.13***	.21***	.17***	.17**	.09***	.10***
ΔR^2		.05**		.05**		.08**		.00		.01*

	Inclusion Climate		Job Satisfaction		Commitment		Incivility		Turnover	
Predictor	Step 1	Step 2	Step 1	Step 2	Step 1	Step 2	Step 1	Step 2	Step 1	Step 2
Nembhard & Edmondson, 2006	.55***	.27**	.56***	.24**	.45***	.06	-.36***	-.27**	-.46***	-.22
ILQ		.38**		.45**		.55**		-.12		-.35*
R^2	.37**	.44**	.28***	.34***	.14***	.21***	.14***	.15***	.08***	.10***
ΔR^2		.07**		.06**		.07**		.01		.02*

Note: ***p < .001, **p < .01, *p < .05, in all cases unstandardized beta coefficients are presented.

Figures

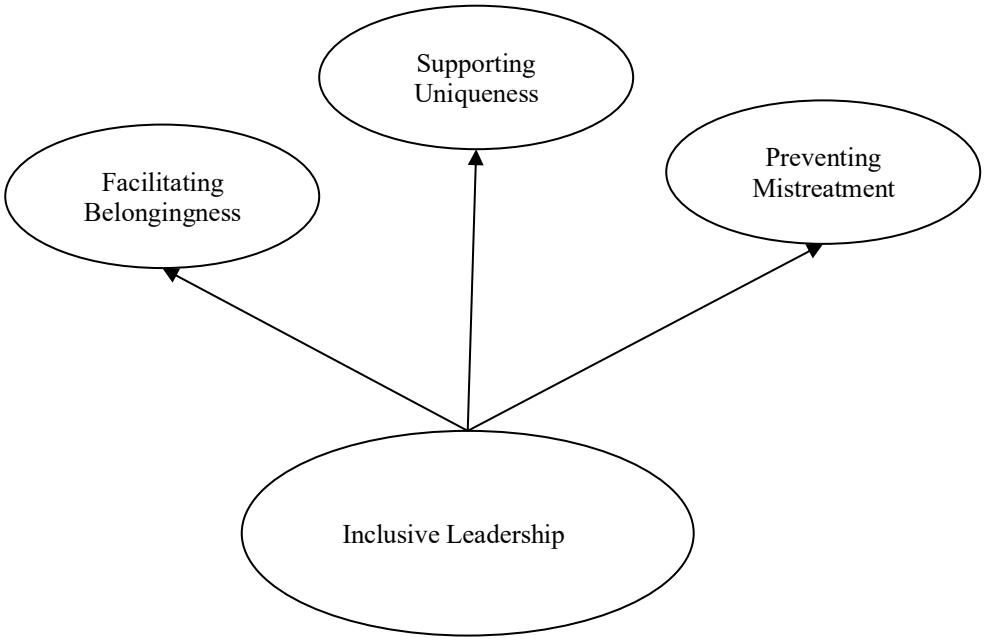

Figure 1: Hierarchical model representing the Inclusive Leadership construct

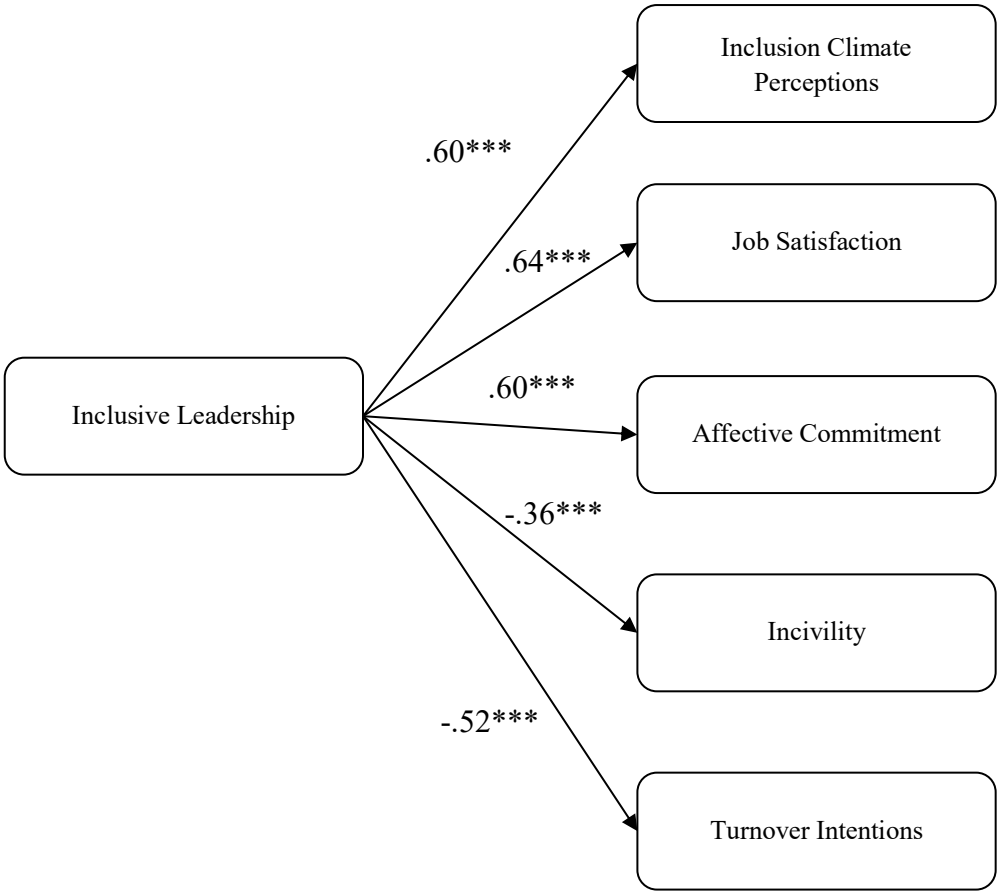

Figure 2: Predictive validity results
****p < .001*

References

Allen, N. J., & Meyer, J. P. (1990). The measurement and antecedents of affective, continuance and normative commitment to the organization. *Journal of Occupational Psychology*, *63*(1), 1–18. https://doi.org/10.1111/j.2044-8325.1990.tb00506.x

Anderson, J. C., & Gerbing, D. W. (1991). Predicting the performance of measures in a confirmatory factor analysis with a pretest assessment of their substantive validities. *Journal of Applied Psychology, 76*(5)*,* 732–740. https://doi.org/10.1037/0021-9010.76.5.732

Antonakis, J., Avolio, B. J., & Sivasubramaniam, N. (2003). Context and leadership: An examination of the nine-factor full-range leadership theory using the Multifactor Leadership Questionnaire. *The Leadership Quarterly*, *14*(3), 261–295. https://doi.org/10.1016/S1048-9843(03)00030-4

Acquavita, S. P., Pittman, J., Gibbons, M., & Castellanos-Brown, K. (2009). Personal and organizational diversity factors' impact on social workers' job satisfaction: Results from a national internet-based survey. *Administration in Social Work*, *33*(2), 151–166. https://doi.org/10.1080/03643100902768824

Ahmed, F., Zhao, F., & Faraz, N. A. (2020). How and when does inclusive leadership curb psychological distress during a crisis? Evidence from the COVID-19 outbreak. *Frontiers in Psychology*, *11*(1898), 1–12. https://doi.org/10.3389/fpsyg.2020.01898

Ashikali, T., & Groeneveld, S. (2015). Diversity management in public organizations and its effect on employees' affective commitment: The role of transformational leadership and the inclusiveness of the organizational culture. *Review of Public Personnel Administration, 35*(2), 146–168. https://doi.org/10.1177/0734371X19899722

Avolio, B. J., & Bass, B. M. (2004). *Multifactor Leadership Questionnaire. Manual and sampler set* (3rd ed.). Mindgarden.

Bae, K. B., Sabharwal, M., Smith, A. E., & Berman, E. (2017). Does demographic dissimilarity matter for perceived inclusion? Evidence from public sector employees. *Review of Public Personnel Administration, 37*(1), 4–22. https://doi.org.10.1177/0734371X16671367

Barling, J., Weber, T., & Kelloway, K. (1996). Effects of transformational leadership on attitudinal and financial outcomes. *Journal of Applied Psychology, 81*(6), 827–832. https://doi.org/10.1037/0021-9010.81.6.827

Bartel-Sheehan, K., & Pittman M. (2016). *Amazon's Mechanical Turk for academics.* Melvin & Leigh.

Bartlett, M. S. (1950). Tests of significance in factor analysis. *British Journal of Statistical Psychology, 3(2),* 77–85. https://doi.org/10.1111/j.2044-8317.1950.tb00285.x

Bass, B., & Avolio, B. (1993). Transformational leadership and organizational culture. *Public Administration Quarterly, 17*(1), 112–121. https://doi.org/10.1080/01900699408524907

Bass, B. M., & Avolio, B. J. (1997). *Full range leadership development manual for the multifactor leadership questionnaire.* Mindgarden.

Bass, B. M., & Bass, R. (2008). *The Bass handbook of leadership: Theory, research, and applications*. Free Press.

Bass, B. M., & Riggio, R. E. (2006). *Transformational leadership* (2nd ed.). Lawrence Erlbaum.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117*(3), 497–529. https://doi.org/0033-2909/95/S3.00

Bilimoria, D., Joy, S. and Liang, X.F. (2008) Breaking barriers and creating inclusiveness: Lessons of organizational transformation to advance women faculty in academic science and engineering. *Human Resource Management, 47(3),* 423–441. http://doi.org/10.1002/hrm.20225

Boekhorst, J. A. (2015). The role of authentic leadership in fostering workplace inclusion: A social information processing perspective. *Human Resource Management, 54*(2), 241–264. https://doi.org/10.1002/hrm.21669

Bonett, D. G., & Wright, T. A. (2000). Sample size requirements for estimating Pearson, Spearman and Kendall correlations. *Psychometrika*, *65*(1), 23–28. https://doi.org/10.1007/BF02294183

Bowling, N. A., & Hammond, G. D. (2008). A meta-analytic examination of the construct validity of the Michigan Organizational Assessment Questionnaire Job

Satisfaction Subscale. *Journal of Vocational Behavior*, *73*(1), 63–77. https://doi.org/10.1016/j.jvb.2008.01.004

Brewer, M. B. (1991). The social self: On being the same and different at the same time. *Personality and Social Psychology Bulletin, 17*(5), 475–482. https://doi.org/10.1177/0146167291175001

Brimhall, K. C. (2019). Inclusion and commitment as key pathways between leadership and nonprofit performance. *Nonprofit Management & Leadership, 30*(1), 31–49. https:// doi.org/10.1002/nml.21368

Brimhall, K. C., Lizano, E. L., & Mor Barak, M. E. (2014). The mediating role of inclusion: A longitudinal study of the effects of leader–member exchange and diversity climate on job satisfaction and intention to leave among child welfare workers. *Children and Youth Services Review*, *40*, 79–88. https://doi.org/10.1016/j.childyouth.2014.03.003

Brimhall, K. C., & Palinkas, L. (2020). Using mixed methods to uncover inclusive leader behaviors: A promising approach for improving employee and organizational outcomes. *Journal of Leadership & Organizational Studies*, *27*(4), 357–375. https://doi.org/10.1177/1548051820936286

Brown, M. E., & Mitchell, M. S. (2010). Ethical and unethical leadership: Exploring new avenues for future research. *Business Ethics Quarterly, 20*(4), 583–616. https://doi.org/10.5840/beq201020439

Brown, M. E., Treviño, L. K., & Harrison, D. A. (2005). Ethical leadership: A social learning perspective for construct development and testing. *Organizational*

Behavior and Human Decision Processes, *97*(2), 117–134. https://doi.org/10.1016/j.obhdp.2005.03.002

Buhrmester, M. D., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *6*(1), 3–5. https://doi.org/10.1177/1745691610393980

Cabrera-Nguyen, P. (2010). Author guidelines for reporting scale development and validation results in the *Journal of the Society for Social Work and Research*. *Journal of the Society for Social Work and Research, 1*(2), 99–103. https://doi.org/10.5243/jsswr.2010.8

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*(2), 81–105. https://doi.org/10.1037/h0046016

Cammann, C., Fichman, M., Jenkins, G. D., & Klesh, J. (1983). Michigan Organizational Assessment Questionnaire. In S. E. Seashore, E. E. Lawler, P. H. Mirvis, & C. Cammann (Eds.), *Assessing organizational change: A guide to methods, measures, and practices* (pp. 71–138). Wiley-Interscience.

Cammann, C., Fichman, M., Jenkins, D., & Klesh, J. (1979). *The Michigan Organizational Assessment Questionnaire* [Unpublished manuscript]. University of Michigan, Ann Arbor.

Carless, S. A., Wearing, A. J., & Mann, L. (2000). A short measure of transformational leadership. *Journal of Business & Psychology, 14*(3), 389–405. https://doi.org/10.1023/A:1022991115523

Carmeli, A., Reiter-Palmon, R., & Ziv, E. (2010). Inclusive leadership and employee involvement in creative tasks in the workplace: The mediating role of psychological safety. *Creativity Research Journal*, *22*(3), 250–260. https://doi.org/10.1080/10400419.2010.504654

Celik, S., Ashikali, T., & Groeneveld, S. (2011). De invloed van diversiteitsmanagement op de binding van werknemers in de publieke sector. De rol van transformationeel leiderschap [The binding effect of diversity management on employees in the Dutch public sector. The role of transformational leadership]. *Tijdschrift voor HRM, 14*(4), 32–53. Retrieved from http://hdl.handle.net/1765/100787

Cho, S., & Mor Barak, M. E. (2008). Understanding of diversity and inclusion in a perceived homogeneous culture: A study of organizational commitment and job performance among Korean employees. *Administration in Social Work, 32*(4)*,* 100–126. https://doi.org/10.1080/03643100802293865

Choi, S. B., Tran, T. B. H., & Park, B. I. (2015). Inclusive leadership and work engagement: Mediating roles of affective organizational commitment and creativity. *Social Behavior and Personality*, *43*(6), 931–943. https://doi.org/10.2224/sbp.2015.43.6.931

Chrobot-Mason, D., & Aramovich, N. P. (2013). The psychological benefits of creating an affirming climate for workplace diversity. *Group & Organization Management*, *38*(6), 659–689. https://doi.org/10.1177/1059601113509835

Chrobot-Mason, D., Ruderman, M. N., & Nishii, L. H. (2014). Leadership in a diverse workplace. In D. V. Day (Ed.), *The Oxford handbook of leadership and organizations* (pp. 683–708). Oxford University Press. http://doi.org/10.1093/oxfordhb/9780199755615.013.034

Chung, B. G., Ehrhart, K. H., Shore, L. M., Randel, A. E., Dean, M. A., & Kedharnath, U. (2020). Work group inclusion: Test of a scale and model. *Group & Organization Management*, *45*(1), 75–102. https://doi.org/10.1177/1059601119839858

Churchill, Gilbert A. (1979). A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research, 16*(1)*,* 64–73. https://doi.org/10.1177/002224377901600110.

Conger J. A., Kanungo R. N. (1987). Toward a behavioral theory of charismatic leadership in organizational settings. *Academy of Management Review, 12*(4), 637–647. https://doi.org/10.2307/258069

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*(1), 98–104. https://doi.org/10.1037/0021-9010.78.1.98

Cortina, L. M., Magley, V. J., Williams, J. H., & Langhout, R. D. (2001). Incivility in the workplace: Incidence and impact. *Journal of Occupational Health Psychology*, *6*(1), 64–80. https://doi.org/10.1037/1076-8998.6.1.64

Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (Neo-PI-R) and NEO Five-Factor Inventory (NEO-FFI): Professional manual*. Psychological Assessment Resources.

Cox, T. H., & Blake, S. (1991). Managing cultural diversity: Implications for organizational competitiveness. *The Executive*, *5*(3), 45–56. https://doi.org/10.5465/ame.1991.4274465

Dalal, R. S. (2013). Job attitudes: Cognition and affect. In N. W. Schmitt & S. Highhouse (Eds.), *Handbook of Psychology Volume 12: Industrial and Organizational Psychology* (pp. 341-366). John Wiley & Sons, Inc.

Deci, E. L., Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior.* Plenum Press.

DeVellis, R. F. (2003). *Scale development: Theory and applications* (2nd ed). Sage Publications, Inc.

Dimoff, J. K., & Kelloway, E. K. (2019). With a little help from my boss: The impact of workplace mental health training on leader behaviors and employee resource utilization. *Journal of Occupational Health Psychology*, *24*(1), 4–19. https://doi.org/10.1037/ocp0000126

Dover, T. L., Major, B., & Kaiser, C. R. (2016). Members of high-status groups are threatened by pro-diversity organizational messages. *Journal of Experimental Social Psychology*, *62(1)*, 58–67. https://doi.org/10.1016/j.jesp.2015.10.006

Edmondson, A.C. (2004). Psychological safety, trust, and learning in organizations: A group-level lens. In R.M. Kramer, & K.S. Cook (Eds.), *Trust and distrust in organizations: Dilemmas and approaches* (pp. 239-272). Russell Sage Foundation.

Ehrhart, K. H., Chung-Herrera, B. G., Randel, A. E., Dean, M. A., & Shore, L. M. (2014, May). Inclusion and health moderated by demographic status as numerical majority/minority [Symposium]. In B. Chung (Chair), *Moving from diversity to inclusion: New directions in inclusion research* [Symposium]. The 29th Annual Conference of the Society for Industrial and Organizational Psychology Conference, Honolulu, HI, United States.

Eisenberger, R., Huntington, R., Hutchison, S., & Sowa, D. (1986). Perceived organizational support. *Journal of Applied Psychology*, *71*(3), 500–507. https://doi.org/10.1037/0021-9010.71.3.500

Eisenberger, R., Stinglhamber, F., Vandenberghe, C., Sucharski, I. L., & Rhoades, L. (2002). Perceived supervisor support: Contributions to perceived organizational support and employee retention. *Journal of Applied Psychology*, *87*(3), 565–573. https://doi.org/10.1037//0021-9010.87.3.565

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*(3)*,* 272–299. https://doi.org/10.1037/1082-989X.4.3.272

Fang, Y.-C., Chen, J.-Y., Wang, M.-J., & Chen, C.-Y. (2019). The impact of inclusive leadership on employees' innovative behaviors: the mediation of psychological capital. *Frontiers in Psychology*, *10* (1803), 1–11. https://doi.org/10.3389/fpsyg.2019.01803

Feitosa, J., Joseph, D. L., & Newman, D. A. (2015). Crowdsourcing and personality measurement equivalence: A warning about countries whose primary language is not English. *Personality and Individual Differences, 75(3),* 47–52. https://doi.org/10.1016/j.paid.2014.11.017

Ferdman, B. M., & Deane, B. R. (Eds.). (2014). *Diversity at work: The practice of inclusion.* Jossey-Bass.

Findler, L., Wind, L. H., & Barak, M. E. M. (2007). The challenge of workforce management in a global society: Modeling the relationship between diversity, inclusion, organizational culture, and employee well-being, job satisfaction and organizational commitment. *Administration in Social Work*, *31*(3), 63–94. https://doi.org/10.1300/J147v31n03_05

Fleischer, A., Mead, A. D., & Huang, J. (2015). Inattentive responding in MTurk and other online samples. *Industrial and Organizational Psychology, 8*(2)*,* 196–202. https://doi.org/10.1017/iop.2015.25

Ford, J. K., MacCallum, R. C., & Tait, M. (1986). The application of exploratory factor analysis in applied psychology: A critical review and analysis. *Personnel Psychology*, *39*(2), 291–314. https://doi.org/10.1111/j.1744-6570.1986.tb00583.x

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50. https://doi.org/10.1177/002224378101800104

Fry, E. (1977). Fry's readability graph: Clarifications, validity, and extension to level 17. *Journal of Reading, 21(3),* 242–252.

Gould-Williams, J., & Davies, F. (2005). Using social exchange theory to predict the effects of HRM practice on employee outcomes. *Public Management Review, 7*(1)*,* 1–24. https://doi.org/10.1080/1471903042000339392

Graen, G. B., & Uhl-Bien, M. (1995). Relationship-based approach to leadership: Development of leader-member exchange (LMX) theory of leadership over 25 years: Applying a multi-level multi-domain perspective. *The Leadership Quarterly*, *6*(2), 219–247. https://doi.org/10.1016/1048-9843(95)90036-5

Greenberg, J. (1987). A taxonomy of organizational justice theories. *Academy of Management Review*, *12*(1), 9–22. https://doi.org/10.5465/amr.1987.4306437

Groysberg, B., & Connolly, K. (2013, September 1). *Great leaders who make the mix work*. Harvard Business Review. https://hbr.org/2013/09/great-leaders-who-make-the-mix-work

Guadagnoli, E., & Velicer, W. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin, 103*(2)*,* 265–275. https://doi.org/10.1037//0033-2909.103.2.265

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2009). *Multivariate data analysis* (7th ed.). Pearson.

Hammer, L. B., Johnson, R. C., Crain, T. L., Bodner, T., Kossek, E. E., Davis, K. D., . . . Berkman, L. (2016). Intervention effects on safety compliance and citizenship behaviors: Evidence from the work, family, and health study. *Journal of Applied Psychology, 101*(2), 190–208. http://doi.org/10.1037/apl0000047

Hammer, L. B., Wan, W. H., Brockwood, K. J., Bodner, T., & Mohr, C. D. (2019). Supervisor support training effects on veteran health and work outcomes in the civilian workplace. *Journal of Applied Psychology*, *104*(1), 52–69. https://doi.org/10.1037/apl0000354

Harrison, D. A., & McLaughlin, M. E. (1993). Cognitive processes in self-report responses: Tests of item context effects in work attitude measures. *Journal of Applied Psychology*, *78*(1), 129–140. https://doi.org/10.1037/0021-9010.78.1.129

Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods, 7*(2), 191–205. https://doi.org/10.1177/1094428104263675

Hebl, M. R., & Avery, D. R. (2012). Diversity in organizations. In I. B. Weiner (Ed.), *Handbook of Psychology, Industrial and Organizational Psychology* (2nd ed., pp. 677–697). John Wiley & Sons, Inc.

Hendrickson, A. E., & White, P. O. (1964). Promax: A quick method for rotation to oblique simple structure. *British Journal of Statistical Psychology, 17*(1)*,* 65–70. https://doi.org/10.1111/j.2044-8317.1964.tb00244.x

Hershcovis, M. S. (2011). "Incivility, social undermining, bullying…oh my!": A call to reconcile constructs within workplace aggression research. *Journal of Organizational Behavior*, *32*(3), 499–519. https://doi.org/10.1002/job.689

Hewlin, P. F. (2009). Wearing the cloak: Antecedents and consequences of creating facades of conformity. *Journal of Applied Psychology*, *94*(3), 727–741. https://doi.org/10.1037/a0015228

Hillygus, D. S., Jackson, N., & Young, M. (2014). Professional respondents in non-probability online panels. In M. Callegaro, R. P. Baker, J. Bethlehem, A. S. Goritz, J. A. Krosnick, & P. J. Lavrakas (Eds.), *Online panel research: A data quality perspective* (Vol. 1, pp. 219–237).

Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods*, *1*(1), 104–121. https://doi.org/10.1177/109442819800100106

Hirak, R., Peng, A. C., Carmeli, A., & Schaubroeck, J. M. (2012). Linking leader inclusiveness to work unit performance: The importance of psychological safety and learning from failures. *The Leadership Quarterly, 23*(1), 107–117.

Hobfoll, S. E. (1989). Conservation of resources: A new attempt at conceptualizing stress. *American Psychologist, 44*(3), 513–524. https:///doi.org/10.1037/0003-066X.44.3.513

Hollander, E. P. (2009). *Inclusive leadership: The essential leader-follower relationship.* Routledge.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*(2)*,* 179–185. https://doi.org/10.1007/BF02289447

House, R. J., & Mitchell, T. R. (1974). Path-goal theory of leadership. *Journal of Contemporary Business, 4*(3)*,* 81–97.

Howard, J., Shaw, E., Felsen, C., & Crabtree, B. (2012). Physicians as inclusive leaders: Insights from a participatory quality improvement intervention. *Quality Management in Health Care*, *21*(3), 135–145. https://doi.org/10.1097/QMH.0b013e31825e876a

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Huang, J., Krivkovich, A., Starikova, I., Yee, L., & Zanoschi, D. (2019). *Women in the Workplace 2019.* McKinsey & Company. https://www.mckinsey.com/featured-insights/gender-equality/women-in-the-workplace-2019

Hunt, V., Yee, L., Prince, S., & Dixon-Fyle, S. (2018). Delivering growth through diversity. McKinsey & Company. https://www.mckinsey.com/business-functions/organization/our-insights/delivering-through-diversity

Hutchison, S. (1997). Perceived organizational support: Further evidence of construct validity. *Educational and Psychological Measurement, 57,* 1025–1034. https://doi.org/10.1177/0013164497057006011

Hwang, J., & Hopkins, K. M. (2012). Organizational inclusion, commitment, and turnover among child welfare workers: A multilevel mediation analysis. *Administration in Social Work, 36*(1), 23–39. https://doi.org/10.1080/03643107.2010.537439

Jackson, S. E., & Joshi, A. (2011). Work team diversity. In S. Zedeck (Ed.), *APA Handbook of Industrial and Organizational Psychology, Vol 1: Building and Developing the Organization* (pp. 652-686). American Psychological Association. http://dx.doi.org/10.1037/12169-020

Jansen, W. S., Otten, S., van der Zee, K. I., & Jans, L. (2014). Inclusion: Conceptualization and measurement. *European Journal of Social Psychology, 44*(4)*,* 370–385. https://doi.org/10.1002/ejsp.2011

Jehn, K. A., & Bezrukova, K. (2010). The faultline activation process and the effects of activated faultlines on coalition formation, conflict, and group outcomes. *Organizational Behavior and Human Decision Processes, 112*(1), 24–42. https://doi.org/10.1016/j.obhdp.2009.11.008

Jehn, K. A., Bezrukova, K., & Thatcher, S. (2008). Conflict, diversity, and faultlines in workgroups. In C. K. W. De Dreu & M. J. Gelfand (Eds.), *The psychology of conflict and conflict management in organizations* (pp. 179-210). Taylor & Francis.

Jin, M., Lee, J., & Lee, M. (2017). Does leadership matter in diversity management? Assessing the relative impact of diversity policy and inclusive leadership in the public sector. *Leadership & Organization Development Journal*, *38*(2), 303–319. https://doi.org/10.1108/LODJ-07-2015-0151

Jones, K. P., Peddie, C. I., Gilrane, V. L., King, E. B., & Gray, A. L. (2016). Not so subtle: A meta-analytic investigation of the correlates of subtle and overt discrimination. *Journal of Management*, *42*(6), 1588–1613. https://doi.org/10.1177/0149206313506466

Kelloway, E. K., Barling, J., & Helleur, J. (2000). Enhancing transformational leadership: The roles of training and feedback. *Leadership & Organization Development Journal, 21*(3), 145–149. https://doi.org/10.1108/01437730010325022

Kerlinger, F. (1986). *Foundations of behavioral research* (3rd ed.). Holt, Rinehart, & Winston.

Kernis, M. H., & Goldman, B. M. (2006). A multi-component conceptualization of authenticity: Theory and research. In M.P. Zanna (Ed.), *Advances in Experimental Social Psychology* (pp. 284–357)*,* Academic Press. http://doi.org/10.1016/S0065-2601(06)38006-9

Kim, J.-O., & Mueller, C. W. (1978). *Factor analysis: Statistical methods and practical issues*. SAGE.

Kim, W. G., Leong, J. K., Lee, Y (2005). Effect of service orientation on job satisfaction, organizational commitment, and intention of leaving in a casual dining chain

restaurant. *Hospitality Management, 24*(2), 171–193. https://doi.org/10.1016/j.ijhm.2004.05.004

King, E. B., Dawson, J. F., Kravitz, D. A., & Gulick, L. M. V. (2012). A multilevel study of the relationships between diversity training, ethnic discrimination and satisfaction in organizations. *Journal of Organizational Behavior*, *33*(1), 5–20. https://doi.org/10.1002/job.728

Kline, R. B. (2010). *Principles and practice of structural equation modeling* (3rd ed.). Guilford Press.

Kottke, J. L., & Sharafinski, C. E. (1988). Measuring perceived supervisory and organizational support. *Educational and Psychological Measurement, 48*(4), 1075–1079. https://doi.org/10.1177/0013164488484024

Leslie, L. M., Bono, J. E., Kim, Y. (S.), & Beaver, G. R. (2020). On melting pots and salad bowls: A meta-analysis of the effects of identity-blind and identity-conscious diversity ideologies. *Journal of Applied Psychology, 105*(5), 453–471. https://doi.org/10.1037/apl0000446

Li, C. R., Lin, C. J., Tien, Y. H., & Chen, C. M. (2015). A multilevel model of team cultural diversity and creativity: The role of climate for inclusion. *The Journal of Creative Behavior, 51*(2), 163–179. http://dx.doi.org/10.1002/jocb.93.

Looney, S. W. (2018). Practical issues in sample size determination for correlation coefficient inference. *SM Journal of Biometrics & Biostatistics*, *3*(1), 1–4. https://doi.org/10.36876/smjbb.1027

Madera, J. M., Dawson, M., & Neal, J. A. (2018). Why investing in diversity management matters: Organizational attraction and person–organization fit. *Journal of Hospitality & Tourism Research*, *42*(6), 931–959. https://doi.org/10.1177/1096348016654973

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. https://doi.org/10.1111/j.1744-6570.2007.00064.x

Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods*, *17*(3), 437–455. https://doi.org/10.1037/a0028085

Meyer, J. P., Stanley, D. J., Herscovitch, L., & Topolnytsky, L. (2002). Affective, continuance, and normative commitment to the organization: A meta-analysis of antecedents, correlates, and consequences. *Journal of Vocational Behavior*, *61*(1), 20–52. https://doi.org/10.1006/jvbe.2001.1842

Mor Barak, M. E. (2005). *Managing diversity: Toward a globally inclusive workplace.* Sage.

Mor Barak, M. E., & Cherin, D. A. (1998). A tool to expand organizational understanding of workforce diversity: Exploring a measure of inclusion-exclusion. *Administration in Social Work, 22*(1), 47–64. https://doi.org/10.1300/J147v22n01_04

Mor Barak, M. E., & Daya, P. (2014). Fostering inclusion from the inside out to create an inclusive workplace. In B. M. Ferdman & B. R. Deane (Eds.), *Diversity at work: The practice of inclusion* (pp. 391–412). Jossey-Bass.

Mor Barak, M. E., Findler, L., & Wind, L. (2001). Diversity, inclusion, and commitment in organizations: International empirical explorations. *Journal of Behavioral and Applied Management*, 2(2), 72–91. Retrieved from https://jbam.scholasticahq.com/article/813-diversity-inclusion-and-commitment-in-organizations-international-empirical-explorations

Mor Barak, M.E., & Levin, A. 2002. Outside of the corporate mainstream and excluded from the work community: A study of diversity, job satisfaction, and well-being. *Journal of Community, Work, & Family, 5*(2), 133–157. http://dx.doi.org10.1080/13668800220146346

Mor Barak, M. E., Levin, A., Nissly, J. A., & Lane, C. J. (2006). Why do they leave? Modeling child welfare workers' turnover intentions. *Children and Youth Services Review, 28*(5)*,* 548–577. http://dx.doi.org/10.1016/j.childyouth.2005.06.003.

Mor Barak, M. E., Lizano, E. L., Kim, A., Duan, L., Rhee, M.-K., Hsiao, H.-Y., & Brimhall, K. C. (2016). The promise of diversity management for climate of inclusion: A state-of-the-art review and meta-analysis. *Human Service Organizations: Management, Leadership & Governance*, *40*(4), 305–333. https://doi.org/10.1080/23303131.2016.1138915

Nembhard, I. M., & Edmondson, A. C. (2006). Making it safe: The effects of leader inclusiveness and professional status on psychological safety and improvement

efforts in health care teams. *Journal of Organizational Behavior*, *27*(7), 941–966. https://doi.org/10.1002/job.413

Netemeyer, R. G., Johnston, M. W., & Burton, S. (1990). Analysis of role conflict and role ambiguity in a structural equation framework. *Journal of Applied Psychology, 75*(2)*,* 148–157. https://doi.org/10.1037/0021-9010.75.2.148

Ng, E., & Stephenson, J. (2016). Individuals, teams, and organizational benefits of managing diversity. In. R. Bendl, I. Bleijenbergh, E. Henttonen, & A. J. Mills (Eds.), *The Oxford handbook of diversity in organizations* (Vol. 1, pp. 235–254). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199679805.013.6

Nishii, L. H. (2013). The benefits of climate for inclusion for gender-diverse groups. *Academy of Management Journal*, *56*(6), 1754–1774. https://doi.org/10.5465/amj.2009.0823

Nishii, L. H., Langevin, A., & Bruyere, S. (2010). *Ageism and the retention of high performers: The positive impact of three forms of inclusion* [Technical report]. SHRM Foundation.

Nishii, L. H., & Mayer, D. M. (2009). Do inclusive leaders help to reduce turnover in diverse groups? The moderating role of leader-member exchange in the diversity to turnover relationship. *Journal of Applied Psychology, 94*(6)*,* 1412–1426. https://doi.org/10.1037/a0017190

Nunally, J. C. (1978). *Psychometric theory*. McGraw-Hill.

Panicker, A., Agrawal, R. K., & Khandelwal, U. (2018). Inclusive workplace and organizational citizenship behavior: Study of a higher education institution, India.

Equality, Diversity and Inclusion: An International Journal, *37*(6), 530–550. https://doi.org/10.1108/EDI-03-2017-0054

Pearson, C. M., & Porath, C. L. (2005). On the nature, consequences and remedies of workplace incivility: No time for "nice"? Think again. *Academy of Management Perspectives*, *19*(1), 7–18. https://doi.org/10.5465/ame.2005.15841946

Peer, E., Vosgerau, J., & Acquisti, A. (2014). Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. *Behavioral Research Methods, 46*(4)*,* 1023–1031. http://doi.org/10.3758/s13428-013-0434-y

Percheski, C. (2008). Opting out? Cohort differences in professional women's employment rates from 1960 to 2005. *American Sociological Review, 73*(3)*,* 497–517. http://doi.org/10.1177/000312240807300307

Perry, E. L., Block, C. J., & Noumair, D. A. (2020). Leading in: Inclusive leadership, inclusive climates and sexual harassment. *Equality, Diversity and Inclusion, Vol. ahead-of-print,* No. ahead-of-print. https://doi.org/10.1108/EDI-04-2019-0120

Peterson, C., & Seligman, M. E. P. (2004). *Character strengths and virtues: A handbook and classification.* Oxford University Press/American Psychological Association.

Pittinsky, T. L. (2010). A two-dimensional model of intergroup leadership: The case of national diversity. *American Psychologist*, *65*(3), 194–200. https://doi.org/10.1037/a0017329

Pless, N., & Maak, T. (2004). Building an inclusive diversity culture: Principles, processes and practice. *Journal of Business Ethics*, *54*(2), 129–147. https://doi.org/10.1007/s10551-004-9465-8

Podsakoff, P. M., MacKenzie S. B., & Podsakoff, N. P (2012) Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology, 63*(1), 539–569. https://doi.org/10.1146/annurev-psych-120710-100452

Porter, L. W., Crampon, W. J., & Smith, F. J. (1976). Organizational commitment and managerial turnover: A longitudinal study. *Organizational Behavior Human Decision Process. 15*(1), 87–98. https://doi.org/10.1016/0030-5073(76)90030-1

Price, J. L., & Mueller, C. W. (1986). *Handbook of organizational measurement*. Pitman Publishing.

Price, J. L., & Mueller, C. W. (1981). A causal model for turnover for nurses. *Academy of Management Journal, 24*(3)*,* 543–565. http://doi.org/10.2307/255574

Ran, S., Liu, M., Marchiondo, L. A., & Huang, J. L. (2015). Difference in response effort across sample types: Perception or reality? *Industrial and Organizational Psychology, 8*(2)*,* 202–208. https://doi.org/10.1017/iop.2015.26

Randel, A. E., Dean, M. A., Ehrhart, K. H., Chung, B. G., & Shore, L. M. (2016). Leader inclusiveness, psychological diversity climate, and helping behaviors. *Journal of Managerial Psychology, 31*(1)*,* 216–234. http://doi.org/10.1108/JMP-04-2013-0123

Randel, A. E., Galvin, B. M., Shore, L. M., Ehrhart, K. H., Chung, B. G., Dean, M. A., & Kedharnath, U. (2018). Inclusive leadership: Realizing positive outcomes through belongingness and being valued for uniqueness. *Human Resource Management Review*, *28*(2), 190–203. https://doi.org/10.1016/j.hrmr.2017.07.002

Ratcliff, N. J., Key-Roberts, M., Simmons, M. J., & Jiménez-Rodríguez, M. (2018). *Inclusive Leadership Survey Item Development*. U.S. Army Research Institute. Retrieved from https://apps.dtic.mil/dtic/tr/fulltext/u2/1060743.pdf

Rhoades, L., Eisenberger, R., & Armeli, S. (2001). Affective commitment to the organization: the contribution of perceived organizational support. *Journal of Applied Psychology, 86*(5)*,* 825–836. https://doi.org/10.1037/0021-9010.86.5.825

Rhoades, L., & Eisenberger, R. (2002). Perceived organizational support: A review of the literature. *Journal of Applied Psychology, 87*(4), 698–714. https://doi.org/10.1037/0021-9010.87.4.698

Roberson, Q. M. (2006). Disentangling the meanings of diversity and inclusion in organizations. *Group & Organization Management*, *31*(2), 212–236. https://doi.org/10.1177/1059601104273064

Rummel, R. J. (1970). *Applied factor analysis.* Northwestern University Press.

Sabat, I. E., Martinez, L. R., & Wessel, J. L. (2013). Neo-activism: Engaging allies in modern workplace discrimination reduction. *Industrial and Organizational Psychology*, *6*(4), 480–485. https://doi.org/10.1111/iops.12089

Schmitt, N., & Stults, D. M. (1986). Methodology review: Analysis of multitrait-multimethod matrices. *Applied Psychological Measurement*, *10*(1), 1–22. https://doi.org/10.1177/014662168601000101

Schneider, B., Ehrhart, M.G., & Macey, W. (2013). Organizational climate and culture. *Annual review of psychology, 64,* 361–388. https://doi.org/10.1146/annurev-psych-113011-143809

Schneider, K. T., Hitlan, R. T., & Radhakrishnan, P. (2000). An examination of the nature and correlates of ethnic harassment experiences in multiple contexts. *Journal of Applied Psychology, 85*(1)*,* 3–12. https://doi.org/10.1037/0021-9010.85.1.3

Schwab, D. P. (1980). Construct validity in organizational behavior. In B. M. Staw & L. L. Cummings (Eds.), *Research in organizational behavior* (Vol. 2, pp. 3–43). JAI.

Schriesheim, C. A., Powers, K. J., Scandura, T. A., Gardiner, C. C., & Lankau, M. J. (1993). Improving construct measurement in management research: Comments and a quantitative approach for assessing the theoretical content adequacy of paper-and-pencil survey-type instruments. *Journal of Management, 19*(2)*,* 385–417. https://doi.org/10.1016/0149-2063(93)90058-U

Shore, L. M., Cleveland, J. N., & Sanchez, D. (2018). Inclusive workplaces: A review and model. *Human Resource Management Review*, *28*(2), 176–189. https://doi.org/10.1016/j.hrmr.2017.07.003

Shore, L. M., Randel, A. E., Chung, B. G., Dean, M. A., Holcombe Ehrhart, K., & Singh, G. (2011). Inclusion and diversity in work groups: A review and model for future research. *Journal of Management*, *37*(4), 1262–1289. https://doi.org/10.1177/0149206310385943

Sigelman, M. & Taylor, J.C. (2021, April 15). To build a diverse company for the long term, develop junior talent. Retrieved from https://www.shrm.org/resourcesandtools/hr-topics/organizational-and-employee-development/pages/develop-junior-talent-.aspx

Smith, N. A., Sabat, I. E., Martinez, L. R., Weaver, K., & Xu, S. (2015). A convenient solution: Using MTurk to sample from hard-to-reach populations. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 8*(2)*,* 220–228. https://doi.org/10.1017/iop.2015.29

Snoeyink, M. J., Martinez, L. R., Hamilton, K. M., & Smith, N. A. (2019, April). Support and advocacy as ally behaviors: The creation and validation of a scale. In Gardner, D. M. & Ryan, A. M. (Chairs), *Allyship: What works and what doesn't* [Symposium]. The Society for Industrial and Organizational Psychology 35th Annual Meeting, Austin, TX, United States.

Snyder, C. R., & Fromkin, H. L. (1980). *Uniqueness: The human pursuit of difference.* Plenum.

Sonnentag, S. (2015). Dynamics of well-being. *Annual Review of Organizational Psychology and Organizational Behavior*, *2*(1), 261–293. https://doi.org/10.1146/annurev-orgpsych-032414-111347

Spector, P. E. (1992). *Summated rating scale construction: An introduction.* Sage.

Spector, P. E., & Jex, S. M. (1998). Development of four self-report measures of job stressors and strain: Interpersonal conflict at work scale, organizational constraints scale, quantitative workload inventory, and physical symptoms inventory. *Journal of Occupational Health Psychology, 3*(4), 356–367. https://doi.org/10.1037/1076-8998.3.4.356

Stone, D. L., & Colella, A. (1996). A model of factors affecting the treatment of disabled individuals in organizations. *The Academy of Management Review*, *21*(2), 352–401. https://doi.org/10.2307/258666

Tabachnick, B. G., & Fidell, L. S. (2012). *Using multivariate statistics* (6th ed.). Pearson.

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & L. W. Austin (Eds.), *Psychology of intergroup relations* (pp. 38–48). Nelson Hall.

Tepper, B. J. (2000). Consequences of abusive supervision. *Academy of Management Journal*, *43*(2), 178–190. https://doi.org/10.5465/1556375

Thomas, R. R. (1990, March 1). *From affirmative action to affirming diversity.* Harvard Business Review. https://hbr.org/1990/03/from-affirmative-action-to-affirming-diversity

U.S. Census Bureau. (2012). *Census Bureau projections show a slower growing, older, more diverse nation a half century from now* (CB12–243). http://www.census.gov/newsroom/releases/archives/population/cb12-243.html

Van Knippenberg, D., De Dreu, C. K. W., & Homan, A. C. (2004). Work group diversity and group performance: An integrative model and research agenda. *Journal of Applied Psychology*, *89*(6), 1008–1022. https://doi.org/10.1037/0021-9010.89.6.1008

Walsh, B. M., Lee, J. (Jessie), Jensen, J. M., McGonagle, A. K., & Samnani, A. K. (2018). Positive leader behaviors and workplace incivility: The mediating role of

perceived norms for respect. *Journal of Business and Psychology*, *33*(4), 495–508. https://doi.org/10.1007/s10869-017-9505-x

Walter, S. L., Seibert, S. E., Goering, D., & O'Boyle, E. H. (2018). A tale of two sample sources: Do results from online panel data and conventional data converge? *Journal of Business and Psychology. 34*(4)*,* 425–452. https://doi.org/10.1007/s10869-018-9552-y

Walumbwa, F. O., Avolio, B. J., Gardner, W. L., Wernsing, T. S., & Peterson, S. J. (2008). Authentic leadership: Development and validation of a theory-based measure. *Journal of Management*, *34*(1), 89–126. https://doi.org/10.1177/0149206307308913

Washington, J., & Evans, N. J. (1991). Becoming an ally. In N. J. Evans, & V. A. Wall (Eds.), *Beyond tolerance: Gays, lesbians and bisexuals on campus* (pp. 195–204). American Association for Counseling and Development.

Wasserman, I. C., Gallegos, P. V., & Ferdman, B. M. 2008. Dancing with resistance: Leadership challenges in fostering a culture of inclusion. In K. M. Thomas (Ed.), *Diversity resistance in organizations* (pp. 175-200). Taylor and Francis.

Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist, 34*(6), 806–838. https://doi.org/10.1177/0011000006288127

Venkatraman, N. 1989. Strategic orientation of business enterprises: The construct, dimensionality, and measurement. *Management Science, 35*(8)*,* 942–962. https://doi.org/10.1287/mnsc.35.8.942

Zhang L., Yang F., & Gu, Y. (2016). Inclusive leadership: Conception, measurement and relationships to related variables. *Advances in Psychological Science*, *24*(9), 1467–1477. https://doi.org/10.3724/SP.J.1042.2016.01467

Zheng, X., Diaz, I., Zheng, X., & Tang, N. (2017). From deep-level similarity to taking charge: The moderating role of face consciousness and managerial competency of inclusion. *Leadership & Organization Development Journal, 38*(1), 89–104. https://doi.org/10.1108/LODJ-06-2015-0134

Appendix A: SME Instruction Sheet

INSTRUCTIONS

The purpose of this scale is to be used as a measure of a leader's inclusiveness when managing diverse teams. For the purposes of this study, a leader refers to anyone who serves as a direct supervisor to other employees. Please write at least three items per dimension, but feel free to write more than that.

I envision some items to specifically focus on preventing discrimination and mistreatment, particularly against employees who come from traditionally marginalized groups. I also intend for the scale to include behaviors that facilitate inclusion for all employees (including majority group members) by fostering belongingness and demonstrating a value for employees' unique perspectives (see definitions of these concepts below). As such, please think about developing items that reflect leader behaviors that foster belongingness and uniqueness more broadly.

Also, please keep in mind the following guidelines when developing items:

- Items should utilize clear and concise statements.
- Items should be written using language that would be considered familiar. Please avoid jargon, expressions, and colloquial language.
- Please write items that reflect behaviors (i.e., actions managers can take) with regard to each of the dimensions.
- Each item should address just one behavior; "double-barreled" items should be avoided.
- Items may overlap/be redundant with one another in terms of general content, but should differ in sentence structure and terminology.
- Avoid leading items. Items should be written in such a way that they are assessed as generally "neutral" so that participants may more freely choose to endorse (or not endorse) specific items.
- Avoid providing negatively-worded (e.g., "not"; "doesn't") items.
- Items should be written at a high school reading level or below.
- Start items with "My supervisor_____."

Finally, the definitions of each dimension are listed below (and in the sheet provided).
Facilitating belongingness: Facilitating employees' human need to be an accepted member of the work group/team and to form and maintain strong, stable interpersonal relationships with others.

Valuing Uniqueness: Supporting employees' need to maintain a distinctive and differentiated sense of self by expressing their unique characteristics and fully contributing to the work group.

Addressing and preventing mistreatment: Ensuring compliance with laws to prevent formal discrimination as well as directly confronting microaggressions to prevent subtle discrimination.

Appendix B: Scale Items

Convergent Validity

Unless otherwise indicated, all scales include responses of 1 to 5: (1) Never, (2) Once in a while, (3) Sometimes, (4) Fairly often, (5) Frequently, if not always.

Inclusive Leadership Questionnaire (ILQ)

Please indicate how often your direct supervisor (manager) behaves in the following ways.

If you have not directly observed a behavior, please rate how frequently you believe or assume they would behave in that way, based on what you know of their leadership style, personal characteristics, or interpersonal relationships.

1. Makes new team members feel included.
2. Makes everyone feel like they belong.
3. Ensures that everyone feels welcome.
4. Asks for different points of view, especially from those who may typically be under- represented.
5. Encourages people to be themselves at work.
6. Allows team members to express their identities however they feel comfortable.
7. Appreciates our team's diversity.
8. Speaks up when someone makes a disrespectful comment regarding marginalized groups.
9. Confronts discrimination when they see it.
10. Stands up for employees who are mistreated.
11. Corrects stereotypical comments people make at work.
12. Creates and/or enforces policies and procedures that reduce bias.
13. Takes appropriate actions to protect the marginalized people in our team.
14. Utilizes feedback from marginalized employees in workplace decisions, policies, and procedures.

Inclusive Leadership (adapted from Nembhard & Edmondson, 2006)

Please indicate how often your direct supervisor (manager) behaves in the following ways.

My supervisor...

1. Encourages all team members to take initiative.

2. Asks for the input of all team members.
3. Values the opinion of all team members equally.

Inclusive Leadership (Carmeli et al., 2010)
Please indicate how often your direct supervisor (manager) behaves in the following ways.

My supervisor...

1. Is open to hearing new ideas.
2. Is attentive to new opportunities to improve work processes.
3. Is open to discuss the desired goals and new ways to achieve them.
4. Is available for consultation on problems.
5. Is an ongoing 'presence' in this team—someone who is readily available.
6. Is available for professional questions I would like to consult with him/her.
7. Is ready to listen to my requests.
8. Encourages me to access him/her on emerging issues.
9. Is accessible for discussing emerging problems.

Perceived Supervisor Support (Eisenberger et al., 1986)
Please indicate your level of agreement with each statement about your immediate supervisor.
Anchors: (1) Agree not at all, (2) Slightly agree, (3) Somewhat agree, (4) Moderately agree, (5) Strongly agree.

My supervisor...

1. Really cares about my well-being.
2. Is willing to extend him/herself in order to help me perform my job to the best of my ability.
3. Cares about my general satisfaction at work.
4. Cares about my opinions.

Transformational Leadership (Carless, 2000)
Please indicate how often your direct supervisor behaves in the following ways.

My supervisor...

1. Communicates a clear and positive vision of the future.
2. Treats staff as individuals and supports and encourages their development.
3. Gives encouragement and recognition to staff.
4. Fosters trust, involvement and cooperation among team members.

5. Encourages thinking about problems in new ways and questions assumptions.
6. Is clear about his/her values and practises what he/she preaches.
7. Instills pride and respect in others and inspires me by being highly competent.

Abusive supervision (Tepper, 2000)
Please indicate how often your direct supervisor (manager) behaves in the following ways.

My supervisor...
1. Makes negative comments about me to others.
2. Gives me the silent treatment.
3. Expresses anger at me when he/she is mad for another reason.
4. Is rude to me.
5. Breaks promises he/she makes.
6. Puts me down in front of others.

Ally Support and Advocacy Behaviors (Snoeyink et al., 2019)
Please indicate your level of agreement with each statement about your immediate supervisor.
Anchors: (1) Agree not at all, (2) Slightly agree, (3) Somewhat agree, (4) Moderately agree, (5) Strongly agree.

My supervisor...
1. Speaks up for a group when nobody who belongs to that group is around.
2. Reminds others to use inclusive language.
3. Confronts someone who uses inappropriate language.
4. Makes suggestions to HR or upper management about changing policies to be more inclusive.
5. Provides a "shoulder to cry on".
6. Listens when someone vents about bias they have experienced.
7. Makes sure someone is OK after they experience prejudice of some kind.

Discriminant Validity

Adventurousness (Adapted from the NEO Personality Inventory (NEO-PI-R); Costa & McCrae, 1992)
To what degree do you agree the following statements describe your supervisor's personality? Anchors: (1) Agree not at all, (2) Slightly agree, (3) Somewhat agree, (4) Moderately agree, (5) Strongly agree.

My supervisor…

1. **Prefers variety to routine.**
2. Likes to visit new places.
3. Is interested in many things.
4. Likes to begin new things.
5. Prefers to try new things.

Humor/Playfulness (Adapted from Values in Action; Peterson & Seligman, 2004)
To what degree do you agree the following statements describe your supervisor's personality? Anchors: (1) Agree not at all, (2) Slightly agree, (3) Somewhat agree, (4) Moderately agree, (5) Strongly agree.

My supervisor…

1. Tries to have fun in all kinds of situations.
2. Tries to add some humor to whatever they do.
3. Has a great sense of humor.
4. Is fun to be with.
5. Goes out of their way to make people laugh.

Orderliness (Adapted from the NEO Personality Inventory (NEO-PI-R); Costa & McCrae, 1992)
To what degree do you agree the following statements describe your supervisor's personality? Anchors: (1) Agree not at all, (2) Slightly agree, (3) Somewhat agree, (4) Moderately agree, (5) Strongly agree.

My supervisor...

1. Likes to keep things tidy.
2. Wants everything to be "just right."
3. Loves order and regularity.
4. Does things according to a plan.
5. Is bothered by disorder.

Predictive Validity

Work Group Inclusion (Chung et al., 2020)

Please indicate the degree to which you personally agree with each of the following statements about the group (team) you belong to at work. Anchors: (1) Agree not at all, (2) Slightly agree, (3) Somewhat agree, (4) Moderately agree, (5) Strongly agree.

1. I am treated as a valued member of my team/work group.
2. I belong in my team/work group.
3. I am connected to my team/work group.
4. People in my team/work group listen to me even when my views are dissimilar.
5. While at work, I am comfortable expressing opinions that diverge from my team/group.
6. I can share a perspective on work issues that is different from my team/group members.

Job satisfaction (Cammann, Fichman, Jenkins, and Klesh, 1983)
Please indicate the degree to which you personally agree with each of the following statements. Anchors: (1) Agree not at all, (2) Slightly agree, (3) Somewhat agree, (4) Moderately agree, (5) Strongly agree.

1. All in all I am satisfied with my job.
2. In general, I like my job.
3. In general, I like working here.

Organizational affective commitment (Allen & Meyer, 1990)
Please indicate your level of agreement with each statement. Anchors: (1) Agree not at all, (2) Slightly agree, (3) Somewhat agree, (4) Moderately agree, (5) Strongly agree.

1. This organization has a great deal of personal meaning for me.
2. I feel 'emotionally attached' to this organization.
3. I feel a strong sense of belonging to my organization.

Turnover intentions (Porter et al., 1976)
Please indicate your level of agreement with this statement. Anchors: (1) Agree not at all, (2) Slightly agree, (3) Somewhat agree, (4) Moderately agree, (5) Strongly agree.

1. I will probably look for a new job during the next year.

Incivility (Cortina, Magley, Williams, and Langhout, 2001)
Please indicate how frequently you have experienced the following types of behaviors from a coworker.

Over the last month, a coworker...

1. Made demeaning or derogatory remarks about me.
2. Interrupted or spoke over me.
3. Ignored me or failed to speak to me ("gave me the silent treatment").
4. Addressed me in unprofessional terms, either publicly or privately.

Incremental Validity

I used the same inclusive leadership measures as in the convergent/discriminant validity phase (Carmeli et al., 2010; Nembhard & Edmondson, 2006) and the same same inclusion climate perceptions measure as in the predictive validity phase (Chung et al., 2020).

Screening Items

These screening items were used in exploratory factor analysis, confirmatory factor analysis, and convergent/discriminant validity stages at the start of the survey to screen out participants who did not meet all inclusion criteria. They were used in the predictive/incremental phase at the end of the survey as part of the demographic survey to identify participants who did not meet inclusion criteria and needed to be removed from analyses.

What is your current employment status?

- Unemployed
- Part-Time Employed
- Full-Time Employed
- Self-Employed

Do you currently work under the supervision of a manager?

- Yes
- No

How many hours per week do you work, on average?

Demographic Items

Please provide the following information about yourself.

What is your current employment status?

- Part-time employed
- Full-time employed
- Self-employed

How many hours per week do you work, on average?

How many of those hours per week do you work remotely?

Which of the following best defines your current gender identity? Select all that apply.

- Female
- Male
- Transgender
- Genderqueer, non-binary, or gender fluid
- Prefer to self-describe: _____
- Prefer not to respond

Which of the following best defines your race or ethnicity? Select all that apply:

- Asian
- Black or African American
- Hispanic, Latino/a/e or Spanish
- Middle Eastern or North African
- Native Hawaiian or other Pacific Islander
- White
- Prefer to self-describe: _____
- Prefer not to respond

What is your age? ___

Do you consider yourself to be:

- Heterosexual or straight
- Gay, lesbian or bisexual
- Prefer to self-describe: _____
- Prefer not to respond

Approximately how many hours per week do you spend directly interacting with your direct supervisor, including time in meetings and online communication tools?

How many years have you worked at your current employer? (If under 1 year, use decimals such as .5 for 6 months).

How many years have you reported to your current supervisor? (If under 1 year, use decimals such as .5 for 6 months).