A paradigm-shift for measuring diversity climate

Gaëlle Cachat-Rosset (contact)

Department of CSR and Human Resources, Toulouse Business School, France

[gaelle.falcon@bbox.fr](mailto:gaelle.falcon@bbox.fr) - +33 (0)6 20 35 56 62

A paradigm-shift for measuring diversity climate

Despite its increasing societal importance, diversity climate (DC) has still remained inconsistently conceptualized and operationalized (Cachat-Rosset et al., 2017; Chrobot-Mason and Aramovich, 2013; Dwertmann et al., 2016; McKay and Avery, 2015). To answer this important issue, this study aims at developing a robust (DC) scale and provide empirical validation for the newly-designed conceptualization from Cachat-Rosset et al. (2017) that consists of three distinct facets: intentionality, programming, praxis. Relying on the *diversity system structure* approach from Arthur and Boyles (2007) and the *diversity management system components* from Kulik (2014), we develop a measure with a structural approach of DC and we question established diversity management paradigms as a way for assessing DC (Dwertmann et al., 2016; Thomas, 1990; Thomas and Ely, 1996). We develop and validate a reliable 11-item scale resulting from a three-stage development process (N=150; N=398; N=221) at individual and higher levels. The nomological validity is established by examining the impact of DC on social integration at unit level. We assess the predictive power of the newly designed instrument by comparing it to the well-established measure from McKay et al. (2008). We finally discuss the implications of our research and propose an innovative taxonomy of organizations’ diversity profiles.

**Keywords:** diversity climate, diversity management, diversity paradigms, multi-level analysis, social integration, scale development and validation

PURPOSE

Modern societies evolve towards what Vertovec (2007) has termed ‘super-diversity’. As far as businesses are concerned, there has been a longstanding support for promoting workforce diversity as it engenders a variety of benefits for organizations and societies (European Commission, 2003). Diversity has been shown to foster innovation (Ozgen et al., 2013; Van der Vegt and Janssen, 2003), creativity and problem solving (Cox and Blake, 1991; Richard et al., 2013) and more diversified customers (Cox, 1993; Herring, 2009; Thomas and Ely, 1996). However, early research suggested that diversity *per se* does not benefit organizations and needs a proper context to blossom (Cox and Blake, 1991). Regardless of contextual considerations, DC has been identified as a key concept expected to strongly moderate the relationship between an organization’s workforce diversity and both individual and organizational positive outcomes, while preventing possible negative ones. It is also expected to be a mediator between human resource (HR) management practices in favor of diversity and the abovementioned outcomes (Cox and Blake, 1991). So, DC stresses the crucial importance of employees’ perceptions for the success of diversity management, especially in a context of increasing signals of resistance towards diversity.

Unfortunately, scholars in the DC body of research have consensually acknowledged that the concept remained poorly conceptualized and operationalized (Cachat-Rosset et al., 2017; Chrobot-Mason and Aramovich, 2013; Dwertmann et al., 2016; McKay and Avery, 2015). Dwertmann et al. (2016, p.1137) argued that conceptual lack of clarity about the DC concept “*have hindered the advancement of diversity climate research*”, ambiguously relying on a *fairness and discrimination* perspective as well as on a *synergy* one and providing confusion in levels of analysis for DC. In a recent attempt to answer the calls for DC concept clarification, Cachat-Rosset et al. (2017) produced a systematic analysis of existing DC literature. They highlighted serious concerns regarding the levels of conceptualization, the fuzzy multi-dimensionality of the concept, inconsistencies between the DC conceptualization and its operationalization and finally an overly dominant Anglo-Saxon context that limits its use in different cultural contexts. Then, relying on organizational climate framework (Schneider and Reichers, 1983), they provided a reconstruction of the concept emerging from the systematic analysis of the conceptual definitions and dimensions of DC and thus reinforcing its conceptual validity. Cachat-Rosset et al. (2017, p.12) defined DC as “*the shared perception by employees of intent, programs, attitudes and behaviors in favor of visible or invisible forms of diversity relevant to the organization’s social context*”. They conceptualized DC as a combination of three perceptual components: intentionality, programming and praxis. Intentionality refers to perceptions that the organization values diversity in discourse such as through formal commitments; the programmatic component refers to the perception of formal diversity policies in the organization; the pragmatic component reflects perceptions about management’s and colleagues’ attitudes and behaviors. Although not questioning the relevance of existing paradigms for diversity management, this disruptive conceptualization of DC challenged them and proposed a new insight in the field with a structural approach, in the sense of Arthur and Boyles (2007). But this innovative way of conceptualizing DC is still not operationalized and needs to be further explored.

We propose in this study to operationalize this new conceptualization of DC by developing a measure instrument aligned with the three-dimension structure proposed by Cachat-Rosset et al. (2017) and by testing it in the DC nomological network. This new approach of DC assessment should allow to appreciate the effectiveness of the organization’s intention regarding diverse people, of the diversity policies to overcome resistance and of the practices implemented at work. To proceed, we assessed the convergent and discriminant validity of the three-factor concept, tested for its incremental validity and investigated its influence on people’s social integration (SI). Our study is intended to extend research in four key ways. First, it provides a needed empirical operationalization of DC which answer the pitfalls highlighted in previous scales. Second, it offers a higher-level DC measure meeting the theoretical background of organizational climate in which the concept is anchored (Schneider and Reichers, 1983). Third, improving the predictive ability of DC measurement to capture higher-level of the DC concept, it will allow the DC stream of research to further explore its influence on proximal as well as distal outcomes, and its moderator or mediator role. Finally, this study proposes a DC measure allowing to address simultaneously several social contexts and to make robust comparisons of cultural boundary conditions for DC and diversity management within and between organizations.

After a presentation of our theoretical background for DC, its linkage with diversity management paradigms, and the existing measurements of DC, we present our scale development strategy and explain how the newly designed instrument was developed and validated at both individual and unit levels. We then test the relationship between this new operationalization of DC and SI for nomological validity. Finally, we discuss the implications of our work for theory and practice and provide directions for future research.

THEORETICAL BACKGROUND

The DC phenomenon has attracted the attention of researchers when they realized that it was a key moderator of the relationship between demographic diversity and its expected beneficial outcomes. DC is also expected to directly influence an array of individual attitudes and behaviors, which in turn impact organizational outcomes (Cox, 1993). DC is thus a key moderator of the diversity-performance relationship (Cox and Blake, 1991; Kossek and Zonia, 1993; Cox, 1993; Mor Barak et al., 1998).

DC at the individual level, also termed psychological DC, has been theorized relying on Cox’s IMCD framework and/or social identity theory (SIT) (Tajfel & Turner, 1986). It is seen as ‘*employees’ individual views of how strongly a work context is viewed as fair and inclusive (or as valuing diversity)*’ (McKay and Avery, 2015, p.196). At the individual level, scholars have demonstrated that positive DC is associated with higher organizational commitment (Buttner et al., 2012; Gonzalez and DeNisi, 2009; Hicks-Clarke and Iles, 2000), job satisfaction (Madera et al., 2013), learning behavior (Guchait et al., 2016), or lower turnover intentions (Buttner et al., 2012; Gonzalez and DeNisi, 2009). DC can also be theorized as an organizational climate (Schneider and Reichers, 1983; Kopelman et al., 1990; Reichers and Schneider, 1990) in relation to the way diverse people are treated and valued by their organization. At higher level, scholars have highlighted the impact of DC on performance (McKay et al., 2008; McKay et al., 2009).

At higher level, DC reflects the employees’ shared perception of the extent to which their organization promotes and positively manages diversity. By managing diversity, organizations seek to develop and strengthen the DC perceived by their employees, which in turn may encourage them to adopt the rewarded behaviors regarding diversity. Dwertmann et al. (2016, p.1137) exposed that DC referred to ‘employees’ perceptions about the extent to which their organization values diversity as evident in the organization’s formal structure, informal values, and social integration of underrepresented employees’. Although the DC concept has been widely defined and used, there have only been limited research efforts to extensively review the DC literature (Cachat-Rosset et al., 2017; Chrobot-Mason and Aramovich, 2013; Dwertmann et al., 2016; McKay and Avery, 2015). However, these reviews agreed on several serious inconsistencies in the DC conceptualization, in the consideration of levels of analysis and in its dimensionality.

We first highlight below the lack of linkage between DC and the diversity management conceptual frameworks. Then, we present the paradigm shift proposed by Cachat-Rosset et al. (2017) to advance the DC field in answer to the call of scholars for clarification. In addition, we outline limitations that pertain to the existing measurement of DC and the need for developing a strongly conceptually aligned new one.

## Grand diversity management paradigms

Diversity management has been conceptualized according several frameworks to further enhance our understanding of what is at stake when organizations manage diversity (Dwertmann et al., 2016; Thomas, 1990; Thomas and Ely, 1996).

The earliest theorist of diversity management, Thomas (1990), distinguished between the Affirmative Action (hiring and promoting persons from under-represented groups) and the Affirming Diversity (valuing their specific contribution) perspectives. In a similar vein, organizations can be categorized as plural (valuing minimal fair treatment and representation) and multi-cultural (valuing the full utilization of diverse employees’ contribution) (Cox, 1991). Then, Thomas and Ely (1996) distinguished between three paradigms: Discrimination & fairness (close to the Affirmative Action approach along the lines of Thomas, 1990), Access & legitimacy, Learning & effectiveness (the last two being related to Thomas (1990)’s Affirming Diversity approach). Discrimination & fairness refers to legal compliance based on efforts to recruit under-represented groups and avoid discrimination. Access & legitimacy refers to tapping diversity among employees in order to better serve diversified customers. Learning & effectiveness refers to using workforce diversity as a leverage for organizational learning. The last two paradigms stress the accrued business benefits and synergy from valuing workforce diversity, while the first paradigm which is highly dominant in DC research, focuses on fair treatment (Dwertmann et al., 2016).

-------------------------------------------

INSERT TABLE I ABOUT HERE

-------------------------------------------

In summary, we fully concur with Dwertmann et al. (2016) who exposed that mainstream diversity management theoretical frameworks can be grouped along a distinction between fair representation and treatment on the one hand, and valuing diversity on the other hand (Table I). However, for the time being DC research most often fails to reflect these distinct diversity management paradigms in their operationalization. Dwertmann et al. (2016) highlighted that DC operationalization is often treated as unidimensional without reflecting the two paradigms under consideration.

## A Diversity Climate paradigm shift

Considering the strong call of scholars for clarifying the DC construct, Cachat-Rosset et al. (2017) proposed a new conceptual model for DC. Following the ‘construct alignment’ principles (Van Campenhoudt and Quivy’s, 2011), they deconstructed the previous DC conceptualizations, definitions, and operationalizations through a systematic analysis of 62 academic articles referring to DC. Their analysis confirmed the multi-dimensionality of DC and allowed them to identify recurrent dimensions addressed when exploring DC.

They derived that DC is the combination of three dimensions: *intentionality*, *programming* and *praxis* (see Table II). Intentionality refers to employees’ perceptions that the organization values diversity; the programming dimension refers to the perception of formal diversity programs or policies and action plans set up within the organization to develop and support workforce diversity; the praxis dimension reflects perceptions about management’s and colleagues’ attitudes and behaviors towards diverse people. DC is then defined as “*the shared perception by employees of intent, programs, attitudes and behaviors in favor of relevant forms of diversity in the social context of the organization*” (Cachat-Rosset et al., 2017). Most previous DC definitions include one or several of these components but almost all of them fail to cover all three dimensions in a combined definition and associated measure. So, their contribution proposes an alternative and disruptive framework for conceptualizing DC at higher level.

-------------------------------------------

INSERT TABLE II ABOUT HERE

-------------------------------------------

By doing this, they proposed a shift of paradigm for DC study, getting ‘free’ from the diversity management paradigms. This new conceptualization proposes to anchor DC research in a diversity management system structure approach (Arthur and Boyles, 2007), where components of the diversity system may be perceived as meaning by employees to shape the organization’s DC. We rely on Arthur and Boyles (2007), who consider that HR climate perception is the shared perception and interpretation of the meaning of HR principles, policies, programs and practices in people’s organization, and we extend it to the DC field. Consistent with the organizational climate literature, Arthur and Boyles (2007) noted that the climate component focuses on employees’ perception of organizational expectations and rewards, and that they derived it from the combination of principles, policies, programs and practices. This inspired Kulik (2014) who applied a levels components typology derived from Arthur and Boyles (2007) to define their ‘*diversity management system components’*. In this five-component system for describing diversity management, DC is the fifth level and is defined as the ‘*shared employee [unit-level] perceptions and interpretations of the meaning of diversity paradigms, policies and programmes in the organization*’ (Kulik, 2014, p.132). In the same vein, Hobman et al. (2004) emphasized that DC is reflected through HR policies and procedures and general attitude toward the value of a diverse workforce for organizational effectiveness. This echoes the call from McKay and Avery (2015) to further explore the strategic HR management literature as a relevant theoretical perspective for diversity management and DC antecedents. Relying on the concept of a ‘strong situation’ (Mischel, 1973), we argued that important messages that shape employees’ perception of organizational expectations regarding diversity are sent by the organization and through the voice of its top management, by proposed formal programs as well as by practices effectively implemented by the proximal actors. And that these messages collectively may create a strong DC when consistently perceived and interpreted by employees (Bowen and Ostroff, 2004). Moreover, DC as any climate results in the subjective reality perceived by employees, whatever the objective reality. So, asking employees for the intention, programs and practices of the organization does not seek to assess if such features objectively exist, but to capture both people ‘awareness’ of these features and their ‘interpretation of the why’, that is the social attribution process (Fiske and Taylor, 1991) that leads them to interpret the organization’s purpose (Nishii et al., 2008).

## Distinctiveness of DC from a major related concept: inclusion climate

In another way and to specify a related but distinct concept, Roberson (2006) distinguishes between diversity and inclusion concepts: diversity focuses on achieving a diverse organizational demography, a notion close to the Affirmative action/Fairness & discrimination perspective (Thomas, 1990; Thomas and Ely, 1996), while inclusion deals with the removal of obstacles to the full contribution of all employees in organizations, closer to the “Affirming Diversity” by Thomas (1990) and the synergy perspective identified by Dwertmann et al. (2016). More recently, Shore et al. (2011) developed a conceptual model of inclusion consisting of two dimensions: perception of *belongingness* and fulfillment of the need for *uniqueness*, clearly distinguishing the concept of inclusion climate from those of DC. Relying on the optimal distinctiveness theory (Brewer, 1991), Shore et al. (2011, p. 1265) defined inclusion in the work group as *‘the degree to which an employee perceives that he or she is an esteemed member of the work group through experiencing treatment that satisfies his or her needs for belongingness and uniqueness’* providing another perspective for inclusion. Chung et al. (2019) operationalized this conceptual model by developing a scale for inclusion assessment at workgroup level, they so confirmed the two-factor structure of inclusion climate. Interestingly, they also demonstrated that DC is an antecedent of workgroup inclusion perception (Chung et al., 2019), confirming the distinctiveness but correlation of the two constructs. On the other hand, Nishii (2013) argued that the construct of climate for inclusion is broader than the DC one. In addition to fair and pro-diversity policies or practices (similarly to DC), climate for inclusion integrates two other dimensions: ‘*integration of differences*’ which captures the interpersonal integration of diverse employees and ‘*inclusion in decision making*’ which refers to the extent to which employees’ diverse perspectives are really integrated (Nishii, 2013).

## Limitations and inconsistences of Diversity Climate measures

Scholars that have attempted to review DC research have raised serious concerns regarding a number of important issues and pitfalls for DC measurement instruments (Table II).

First the “fair treatment and representation” versus “valuing diversity” divide identified in the previous section is poorly reflected in DC operationalizations, since existing DC measures are not clearly designed to tap these distinctions when assessing people’ perception of diversity management efforts. Dwertmann et al. (2016) contend that most DC measures reflect the “fairness & discrimination” perspective and additionally, they report issues in categorizing measures binarily along the “fairness & discrimination” and “synergy” distinction as they find some items to fall somewhere in between the two categories.

Second, DC measurements were sometimes based on demographics data (and not perceptions), in contradiction with the founding assumption that an organizational climate is rooted in perceptions and interpretations drawn by members of the organization about what it does (Katz and Kahn, 1978). Reichers and Schneider (1990, p. 22) defined an organizational climate as the ‘*shared perception of the way things are around here*’, which inherently anchors climate measurement in the perception of individuals as the foci of data collection.

Third, some DC measurement instruments were found to be misaligned with their associated conceptualization, causing concern about the construct’s reliability and validity at different levels of analysis (Chrobot-Mason and Aramovich, 2013; Dwertmann et al., 2016). Especially, there has been a dominance of psychological DC research whereas most scholars used the organizational climate framework to theoretically ground their work. For example, although Mor Barak et al. (1998) grounded their DC definition within the organizational climate conceptual framework (Reichers and Schneider, 1990), they chose to develop a measure that includes dimensions such as personal diversity value and personal comfort with diversity. But an individual’s value towards diversity or personal comfort with diversity may be at odds with his/her organizational DC perception. So, while the authors claimed to introduce an organizational-level construct, their measure is based on individual-referent perceptions. In another case, some measures are not aligned with the proposed definition of DC. For another example, McKay et al. (2008) referred to the integration of under-represented people in their definition of DC, but none of the four items of their scale specifically referred to under-represented people.

Next, there has been a lack of consistency between measures in terms of their respective targets and foci (Cachat-Rosset et al., 2017). In some items, respondents were asked about their individual perceptions about themselves such as “My coworkers appreciate my background and perspective” (Chung et al., 2015, p.1503), in others the emphasis is placed on the respondents’ perceptions about designated groups such as employees of color (Hopkins et al., 2001; Kossek and Zonia, 1993; Kossek et al., 2003), while yet other measures address respondents’ perceptions without specifying an identity group (Chung et al., 2015). It is noteworthy that the mean of individuals’ perception about themselves as an index of a higher-level phenomenon is ungrounded, because the aggregate score would always refer to the individual level and not to the unit or organization level (Guzzo et al., 1993). The referent-shift nature of DC implies that individuals are asked about their perception about a nested level and not about their perception for themselves, to then form a higher-level DC measure (Chan, 1998). Finally, adding to such lack of clarity, the targeted objects in DC measures also vary according to studies. Targeted objects can be policies or processes, employee attitudes or behaviors such as listening, valuing, integrating, while yet others focus on top management commitments/intentions (Dwertmann et al., 2016). Unfortunately, the use of several objects’ perception in the assessment of DC, while it may be suitable, is most often poorly or not conceptually justified regarding the choice of the different objects assessed. Most previous DC definitions include one or several of these components (intentions, programs, attitudes & behaviors) but almost all of them fail to cover all three dimensions in a combined definition and associated measure.

-------------------------------------------

INSERT TABLE III ABOUT HERE

-------------------------------------------

A relevant DC measure should remedy the limitations exposed above: it needs to present a sound alignment with its conceptual definition, be based on perceptions and not on objective data and must not address the respondent’s perceptions towards her-/him-self. Besides, it should not focus on identified groups that are specific to a given country. Therefore, it is crucial to develop higher-level measures if we are to study DC’s role at the organizational level.

In the next section, we present the three-stage development and validation of a new measure instrument echoing the new conceptualization of DC from Cachat-Rosset et al. (2017), aligned with its definition and encompassing its three dimensions. While this promising conceptualization of DC integrates the latest conceptual advancements in DC research and addresses most of the limitations having been raised in current and past DC research, it remains to be empirically validated.

INSTRUMENT DEVELOPMENT AND VALIDATION

Relying upon multiple sources and methods is recommended when developing measurement instruments (MacKenzie et al., 2011; Moore and Benbasat, 1991). Best practices and recommendations were reviewed with the aim of generating a highly valid and reliable measurement instrument (Churchill, 1979; MacKenzie et al., 2011; Moore and Benbasat, 1991), resulting in the implementation of a rigorous multi-stage development strategy. The first stage consisted in generating an initial pool of items and purifying it (Stage 1). Since DC is a specific instance of organizational climate, it must thus be considered as a perceptual construct whose locus of assessment is the individual perception. Such perceptions must then be aggregated in order to capture the higher-level construct under consideration. Consequently, we first validated the developed instrument at the individual level (Stage 2) and then extended its validation to the aggregate level after considering the degree of consensus between respondents to justify the aggregation (Stage 3). The final validation step addressed nomological validity, assessing the impact of DC on perceptions of SI. Table III summarizes the successive steps that were implemented.

-------------------------------------------

INSERT TABLE IV ABOUT HERE

-------------------------------------------

## Stage 1: Scale development

Prior to the development of an initial pool of items, it was essential to clearly determine the nature and dimensionality of the DC construct. The item generation stage started with the generation of an initial pool of candidate items derived from an extensive review of the literature. Then, a card sorting procedure was used with the purpose of identifying the items exhibiting suitable content validity and deleting problematic ones. Finally, the outcome of the item selection stage was reviewed by an expert panel to assess and reinforce its content and discriminant validity.

### Composition model and dimensionality

As previously underscored by Chrobot-Mason and Aramovich (2013), McKay and Avery (2015), Dwertmann et al. (2016), Cachat-Rosset et al. (2017), DC remains poorly defined as a higher-order multidimensional construct. The nature of the relationships between the DC concept and its dimensions must be further investigated. Law et al. (1998) clarified the different natures of second-order constructs: some are *superordinate models*, where latent variables represent a higher abstraction of the reflective sub-dimensions manifested through them, and some are *aggregate* or *profile models*, where variables are a composite of the formative dimensions that compose it. We consider DC as a superordinate model. This means that intentionality, programming and praxis capture people’s perception of the organizational DC and that DC represents the shared variance of its three dimensions. In other words, intentionality, programming and praxis are each one a manifestation of the organizational DC perceptible by organizational incumbents.

Second, as regards a higher-level concept, it is necessary to define its composition model prior to conduct scale development and test hypotheses. Composition models allow to specify the relationships between constructs at different levels of analysis, to identify the locus of data collection and the aggregate nature of constructs (Chan, 1998). DC is portrayed in this study as a *referent-shift consensus model* according to the Chan’s typology. This means that people perceptions examined (the locus of data collection) must address a collective level in their assessment (the group, the organization) to form a higher-level manifestation of the construct when aggregated. Within-group agreement is a necessary condition for justifying aggregation of individual perceptions (Chan, 1998; Glick, 1985).

Following these assumptions, the aggregate level under consideration when investigating DC is a key point to be addressed. In this study, we focus on the unit-level, for aggregate-level manifestation of DC (Beus et al., 2017). Measuring perceived DC at the unit-level is relevant since it is people’s immediate work environment (Chung et al., 2019) and they are exposed to the same organizational intention and programs, but may vary in their experience of their implementation, so they are supposed to be able to assess both the pro-diversity organizational features and proximal practices.

### Item pool generation

The purpose of this step was to generate a pool of items with high content validity. A systematic review of the DC literature resulted in the identification of an initial pool of 34 empirical academic articles published between 1993 and 2016, thanks to the two key words ‘diversity climate’ and ‘climate for diversity’. Within the identified scales, 7 articles had reused existing scales without changes, 11 had adapted previously used scales, and 16 had proposed and developed new scales. The process resulted in the generation of a pool of 114 distinct items.

### Item selection

This stage aimed at reducing the size of the item pool by only retaining the most relevant ones. A preliminary convergent and discriminant validity assessment was conducted through two rounds of a card sorting procedure, an ‘open’ one and a ‘closed’ one (Moore and Benbasat, 1991). For the closed procedure (Anderson and Gerbing, 1991), three scholars experienced in the diversity field were asked to assign each item to its intended dimension (intentionality, programmatic or pragmatic dimension) with the purpose of identifying items that consistently fell under one dimension with a high level of inter-judge agreement.

-------------------------------------------

INSERT TABLE V ABOUT HERE

-------------------------------------------

Table V summarizes the degree of agreement among the three judges during the closed card sorting round. The kappa coefficient is of 0.668 (p < 0.001), signaling a substantial agreement between raters. After several rounds, the judges decided to discard ambiguous and/or redundant items, those that did not capture DC but some other related construct (e.g. inclusion climate, perceived organizational support), those that did not measure DC as conceptually defined (e.g. individual-referent perceptions, personal comfort with diversity), and double-barreled ones.

For the item selection, we also consider the three dimensions of the DC concept under study. The shape of the organizational DC is allowed by people’s shared perceptions and interpretations of signals sent by the organization regarding diversity (Spence, 1974). So, the considered signals may reflect the three dimensions of intentionality, programs and practices. This resulted in selecting items that address (a) the diversity value and commitment of the organization, including through the voice of its top management that conveys the people management strategy and practices rewarded by the organization (Arthur and Boyles, 2007), (b) the existence of formal programs and policies in favor of diversity which translate the intention into actions (Reinwald et al, 2018), (c) the perception of how the intention and programs are implemented by the managers and employees, since inconsistent implementation of diversity policies by managers and employees at the unit level may cause disagreement in employees' DC perceptions.

Finally, we ensured that no item addressed a predefined group, such as gender or race, in line with the Cachat-Rosset et al. (2017)’s definition which state that DC should be appraised as ‘*in favor of visible or invisible forms of diversity relevant to the organization’s social context*’. Furthermore, focusing on specific groups would have made our measure less internationally versatile. This item selection stage resulted in a pool of 26 items: 9 items addressing intentionality, 8 items programs and 9 praxis.

### Item purification

We followed Straub’s recommendations (1989) regarding the qualitative assessment of content validity when designing new measurement scales. To proceed, we conducted a first round of purification through the help of an expert panel. The purpose of the procedure was to verify whether each scale captured the entire domain of its associated dimension. Four experienced scholars well-versed in scale development were asked to review the three developed scales. Each expert received an email containing detailed instructions, the definition of DC along with its three dimensions and associated items. The experts identified 8 items as being unclear or poorly worded, thus needing rephrasing. No single item was identified as being highly problematic. The item development phase thus ended with a pool of 26 refined items.

## Stage 2: Scale validation at individual level

The further refining and initial validation of the newly-developed measurement was conducted at the individual level. Two subsequent quantitative research components were conducted. A first online survey (Study 1), involving a multi-organization sample (N = 150), allowed to further purify the item pool while verifying the dimensionality of the construct, and assessing the initial consistency of the resulting measurement. A second online survey (Study 2) with a multi-organization sample of 398 respondents was conducted to further confirm the quality of the developed measurement and assess model fit.

### Further item purification and dimensionality assessment

This first stage aimed at further purifying the 26 candidate items and at assessing the scale reliability, indicator reliability, convergent and discriminant validity. An Exploratory Factor Analysis (EFA) was also conducted to help purify the items and assess the initial dimensionality of the construct.

For study 1, an online survey was conducted in which respondents were asked to rate the 26 items on a 5-point Likert scale (from 1 “Strongly disagree” to 5 “Strongly agree”). Items were randomly distributed in the survey questionnaire to limit any potential halo effect. We relied on professional social networks to invite individuals to take part in the survey. The sample resulted in 150 respondents. 62% were female and 38% were male; the average age was 30.03 years; 63.31% were employees, managers or executives whereas 36.79% were master’s level students or persons looking for a job. In such case, respondents were asked to think about their former employer for the sake of answering the survey.

First, we examined the inter-item correlations to ensure that items sufficiently correlated with their respective dimension (surpassing the 0.60 threshold) while having low correlations with items from the other dimensions (cross loading gaps above 0.40) (Hair et al., 2011). Then, we assessed commonality (average variance extracted) and reliability (composite reliability and Cronbach’s alpha coefficients). Before dropping any item, we controlled that any manipulation would not affect the content validity of the related dimension and ensured that this resulted in greater scale reliability.

To explore the underlying factor structure and examine the overall dimensionality of the DC construct, we performed an EFA (IBM SPSS version 20), with oblique rotation (Oblimin) considering the three dimensions theoretically capture the whole construct of DC. To proceed, we first confirmed the legitimacy of performing such analysis with the collected dataset (KMO = 0.793; Bartlett’s Test of Sphericity sig.< 0,000 – Field, 2013). In the first round, five components emerged. Several item deletion strategies were followed, discarding items one by one (or in groups). Items were removed based on their factor loadings and cross-loadings. We made sure that the remaining items properly covered the content domain of each dimension. All strategies converged towards the same stable structure, extracting three components (cumulatively explaining 63% of the total variance). The three components did reflect the three conceptual dimensions of the DC proposed by Cachat-Rosset et al. (2017). The refined measurement instrument resulted in 12 items with four items for each dimension.

As a prerequisite for the evaluation of first-order constructs, we first tested for block uni-dimensionality and reliability using three tests (see Table VI), Cronbach’s α, CR and Dillon-Goldstein’s ρ, which all exceeded the 0.7 threshold (Kline, 2013; Nunnally, 1978; Schmiedel et al., 2014). To assess convergent validity, we examined the factor loadings and the average variance extracted (AVE) from each ﬁrst-order construct. The AVE for each DC dimension ranged from 0.579 to 0.674 (Table VI), all above the commonly recommended 0.50 threshold. The factor loadings ranged from 0.636 to 0.873, exceeding the cut-off value of 0.6 (see Table VII). The criteria ensuring convergent validity were thus all satisfactorily met.

-------------------------------------------

INSERT TABLE VI ABOUT HERE

-------------------------------------------

-------------------------------------------

INSERT TABLE VII ABOUT HERE

-------------------------------------------

To assess discriminant validity, we used three criteria: item cross loadings, the Fornell-Larcker criterion (Fornell and Larcker, 1981) and the Heterotrait-Monotrait Ratio of Correlations (HTMT). Table VII shows that items’ loadings on their intended first-order construct are all greater than their cross-loadings with the other dimensions. All items had cross-loading gaps higher than 0.3 except for ‘Clearly, diversity is not important to this organization’ (INT 3) with a loading difference of 0.262 with the programmatic dimension (Table VII). The factor correlation matrix shows that the AVE for each construct is well above the squared correlation between the construct and any other construct (Table VIII). The HTMT test ensures that a reflective construct has the strongest relationships with its own indicators (Hair et al., 2017). All HTMT ratios did not exceed the recommended 0.85 threshold (Table VIII), thus providing extra confidence in terms of discriminant validity.

-------------------------------------------

INSERT TABLE VIII ABOUT HERE

-------------------------------------------

In conclusion, Study 1 allowed to reduce the pool of items from 26 to 12 items while the results demonstrated satisfactory results in terms of consistency, reliability, convergent and discriminant validity.

### Validity assessment

To further increase the robustness of the developed measurement instrument, the superiority of the proposed operationalization and dimensionality of the DC concept had also to be investigated through a fit comparison with alternative DC models (Study 2). To proceed, an online survey was conducted using the 12-item measurement instrument that had been developed and validated at that stage of the research process. The final multi-organization sample consisted of 398 respondents, composed of females for 46%, males for 54%. All were employed in private companies and 38% held a management position. Moreover, to ensure the stability of the measurement in cross-national contexts, the multi-national nature of the sample was ensured with respondents from 30 different countries mainly in Europe, North America, South America and Australia.

The assessment of item reliability coupled with a confirmatory factor analysis revealed some problematic issue regarding the Intentionality dimension. One item, INT3, which was a reversed item, appeared to significantly decrease the Cronbach’s alpha of the Intentionality dimension (α=0.54) while its factor loading was surprisingly low (-0.120\*\*). It was then decided to drop the item. Internal consistency, scale reliability, indicator reliability, convergent validity and discriminant validity were assessed following the same procedures and thresholds presented in the analysis of Study 1 data. All tests were fully satisfactory (See Tables IX, X and XI).

-------------------------------------------

INSERT TABLE IX ABOUT HERE

-------------------------------------------

-------------------------------------------

INSERT TABLE X ABOUT HERE

-------------------------------------------

-------------------------------------------

INSERT TABLE XI ABOUT HERE

-------------------------------------------

Validity assessment could not go without further investigating the superiority of the proposed model in comparison with alternative models. Figure 1 depicts the two models for which rigorous fit evaluations were performed. While Model A considered DC as a second-order reflexive construct (our proposed model), Model B consisted of considering DC as a first-order model with direct reflective indicators

-------------------------------------------

INSERT FIGURE 1 ABOUT HERE

-------------------------------------------

Of the two models, Model A demonstrated the best fit indices meeting the general cut-off values (Hooper et al., 2008) for absolute, incremental and parsimony fit indices (χ²[398] = 167.0, p < 0.000; CMIN/df = 4.07; CFI = 0.943; SRMR = 0.072; RMSEA = 0.088). While Model B had bad results with fit indices being far from the recommended thresholds (χ²[398] = 529.9, p < 0.000; CMIN/df = 12.52; CFI = 0.782; SRMR = 0.112; RMSEA = 0.170). These results clearly demonstrated the superiority of conceptualizing and operationalizing DC as a second-order construct instead of a first-order one, providing additional evidence of the proposed approach’s relevance. Overall, Stage 2 results provided strong support to the psychometric properties of the developed measurement instrument at the individual level. Since DC is here treated as an instance of organizational climate, it was key to pursue our analysis at the aggregate level to assess the validity of the developed scale.

## Stage 3: Scale validation at aggregate level

Data was collected in two organizations through an online survey (Study 3). The first sub-sample consisted in 101 respondents from 5 different subsidiaries of a European manufacturing company. Respondents were 63% female, 36% male, 1% “other”, their average age was 41.7 years and 35% held a management position. The second sub-sample gathered 120 responses from a group composed of 6 distinct hotels. It was 43% female, 54% male, 3%, “other”, the average age was 31.6 years and 25% held a management position. The subsidiary was chosen as the aggregate level for DC assessment because of two key contextual characteristics. First, each subsidiary is relatively independent for implementing HR and diversity policies. Second, people mainly experience practices and integrate socially at the proximal level (Beus et al., 2017). Before proceeding to aggregate level analyses, we confirmed that aggregate data adequately represent the unit’s DC. First, we tested for both within-unit agreement and between-unit variance through the ICC1 and ICC2 criteria. We found low variance within units, high variance between units and a good reliability of the group means (ICC1 = 0.18; ICC2 = 0.92). This made us confident that unit-level DC is well explained by subsidiary membership (LeBreton and Senter, 2008). Second, we assessed within-unit agreement through the rwg(j) criteria to identify whether DC may represent a unit-level manifestation of the construct with enough relevance. We assessed the median rwg(j) with a uniform distribution. It was of 0.95 (ranged from 0.92 to 0.98), indicating a high level of consensus within units.

### Nomological validity assessment

The assessment of the nomological network surrounding a focal construct is an essential step during the scale development process. In order to evaluate the nomological validity of a focal construct, data for constructs that are theoretically connected should be collected (MacKenzie et al., 2011).

**Social integration**. SI deals with the development of relationships with colleagues and peers within the organization and is key to collecting information and adjusting to the work environment (Morrison, 1993). When developing interactions and relationships with colleagues, employees increase their understanding of the work environment and strengthen their feeling of being part of the unit. DC highlights that diverse people are recognized and valued whatever their group identity. Developing diverse relationships with diverse people, employees answer the call of the organization to adopt the required behaviors towards diversity (Reichers and Schneider, 1990). Positive DC reflect a perception that social interactions respect and value differences (Reinwald et al., 2018). People who perceive a positive DC are more likely to develop relationships with all their peers and feel socially integrated, even if they feel different from their colleagues. SI ensures that people feel part of the whole organization and not only of their identity group (Bauer and Green, 1998). Therefore, we expect that a positive DC perception at the unit level will foster stronger perceptions of SI.

As a consequence, we conducted a unit-level analysis of the DC-SI relationship. In this study, SI was assessed using the 4-item scale of Wanberg and Kammeyer-Mueller (2000). All items were rated on a 5-point Likert scale from 1 “Strongly disagree” to 5 “Strongly agree”. As expected, the scale demonstrated appropriate psychometric properties with a Cronbach’α of 0.85 and inter-item correlations being all above 0.5. Gender and Managerial Position were also captured in the survey as control variables. Descriptive statistics as well as correlation coefficients are summarized in Table XII.

-------------------------------------------

INSERT TABLE XII ABOUT HERE

-------------------------------------------

To test the nomological validity of the scale, we performed a hierarchical linear modeling (HLM) analysis (see Table XIII), considering the hierarchical nested structure of our data within units and the induced correlation of errors terms (Klein and Kozlowski, 2000) and because HLM allows to assess multi-level relationships (Hox, 2010).

We first tested a null model with no predictors but considering SI as the dependent variable (model 1, or null model). The results indicated that SI varies significantly across units, since the proportion of variance between units is approximatively 6 % (ICC = 0.06). Then, we assessed a second model with unit-DC as the independent variable at level 2 (model 2). DC unit score was mean-centered to avoid multicollinearity. The deviance statistics difference between the null model (model 1) and the model 2 (∆Dev = -38.2) demonstrated the significance of the incremental variance added by unit-level DC as a predictive variable in Model 2. Furthermore, Model 2 was found to explain an additional 23% of the variance in SI (Δ R²=0.232) providing extra support regarding the effect of unit-level DC. Finally, unit-level DC was found to be positively related to individuals’ SI perceptions (b = 0.570, p < 0.000), supporting our hypothesis.

-------------------------------------------

INSERT TABLE XIII ABOUT HERE

-------------------------------------------

We also ran HLM analyses including the control variables as independent variables. Management position showed no significant relationship with SI, whereas Gender was found to have a positive effect on SI.

### Unit-level model fit assessment

Complementary analyses were conducted in order to perform a DC-SI model fit assessment at the unit-level by comparing the model fit statistics provided for Model A and Model B (see Figure 1 for the model specification and Table XIV for the results). The confirmatory factor analyses confirmed the satisfactory fit level of Model A that had been previously found at the individual level.

-------------------------------------------

INSERT TABLE XIV ABOUT HERE

-------------------------------------------

### Predictive validity

In order to assess the predictive validity of the newly designed scale, it was decided to compare our results with the well-established DC measure from McKay et al. (2008). We used a sub-sample of Study 3 corresponding to the 120 responses gathered from the hotels group. In Table XIII, Model 3 corresponds to the DC-SI relationship using the unit-level DC measure derived from McKay et al. (2008)’s scale. The HLM analysis revealed that the alternative measure was positively related to SI with high significance but had a lower predictive power than our newly-designed one (b = 0.428\*\*\*). These results testify to the predictive validity of our new three-dimensional DC measure.

In conclusion, our rigorous three-stage methodology led to the development of a highly valid and robust DC measure that displays strong psychometric properties both at the individual and at the aggregate levels.

DISCUSSION AND IMPLICATIONS

# Contributions to research and practice

This research has significant implications for both research and practice. First, our measure addresses the requisites that were identified from the scales limitations analysis of the DC literature. It is aligned with the definition and dimensionality proposed by Cachat-Rosset et al. (2017) concept. It is not focused on the respondent him- or herself but on external (to him or her) targets. It does not focus on identity groups that are specific to one country and can therefore be used in cross-national research. It addresses the higher-level of analysis in the sense that it is designed to be used as an aggregate measure. It also targets organizational level foci: top management commitment/stated intentions, organizational policies, attitudes and behaviours manifested in the unit environment, and so avoids personal values or personal comfort perceptions. From a methodological standpoint, our findings support the nomological validity of our proposed measure, providing empirical support to the positive influence of unit-level DC on employees’ perception of SI, so calling for further research on other DC’s proximal and distal outcomes.

Our measurement development contributes to the debate opened by Dwertmann et al. (2016), who claimed that DC measures most often failed to cover the distinct paradigms of fairness/discrimination and synergy. Indeed, these authors have raised serious concern about how the synergy paradigm is at present neglected in existing DC measures. Our new measure both taps the two paradigms but also questions the assumed dimensionality of these paradigms for DC study. Indeed, we found that the three DC facets combine both perspectives identified by Dwertmann et al. (2016), while most previous measures have failed to do so. The praxis dimension for instance combines both the ‘fairness & discrimination perspective’ (PRAX1 and PRAX2) and the synergy perspective (PRAX3 and PRAX4), while the programmatic dimension includes an item that clearly matches the ‘fairness & discrimination perspective’ (PROG3). Therefore, we note that the ‘fairness & discrimination perspective’ is empirically present in two separate dimensions of the newly-designed DC scale (via PROG3 and PRAX1 and PRAX2).

In short, our study raises questions about the empirical robustness of established diversity management paradigms for assessing DC. We contend that these dichotomies might be inherently problematic as they may prove to be somehow inapplicable in real organizational settings. The fairness & discrimination versus synergy dichotomy for instance assumes that on the one hand, an organization may be perceived by its members as providing equal treatment and being fair, while on the other hand, it may be perceived as not valuing each and every worker’s specific contribution to organizational performance and learning on account of their group identity. But if employees perceive that some among them are undervalued, under-listened to and under-considered due to attitudes and behaviours manifested by their managers and/or colleagues on account of their group identity, then we contend that they may perceive this as discrimination, be it indirect or systemic (Agocs, 2002), and not as equal treatment and fairness. This also suggests that empirically disentangling the fairness & discrimination and the synergy perspectives in DC measures may be a dead-end. Inducing top executives in believing that valuing diversity is different from treating diverse employees fairly may actually be detrimental, as they may lose sight of one the fundamentals of diversity management: the systemic nature of discriminatory and exclusionary processes. “Not valuing” ideas and learning opportunities emanating from entire identity groups amounts to treating these groups unfairly.

In this study, we adopted a system structure approach for DC, by extending the HR system structure approach from Arthur and Boyles (2007) to the diversity field and the diversity management system from Kulik (2014) to operationalize the DC concept. Our proposed operationalization allows a view different from the above-mentioned paradigms considering that DC may be captured in a structural approach, while reflecting both paradigms but not being ‘structured’ by them.

This study also has implications for practitioners. Our proposed measurement may become a tool for cross-national research and practice as it is free from country-specific categories that hamper utilization of many existing scales across borders. Most of all, the new measurement represents a concrete tool for practitioners to assess overall DC and its composition. The structural approach adopted will allow practitioners to better answer to this crucial question: ‘How to develop and maintain a positive DC?’. Assessing which of the three dimensions are the most contributive in their DC will provide them with strong guidelines on how to act regarding their diversity management system. Our results suggest that employees may perceive a range of formal and informal signals about diversity management, regarding the intent of their organization to support of diversity, regarding formal policies or procedures to promote diverse people, or regarding attitudes and behaviors of managers or colleagues towards rejecting discrimination and valuing people who are different. While the three DC components taken together form DC perceptions, each one can be more or less prevalent in specific organizations and have separate outcomes. In other words, organizations can be characterized by different DC configurations (see Table XV) outlined in the next section.

-------------------------------------------

INSERT TABLE XV ABOUT HERE

-------------------------------------------

In Type I configuration, an organization may not be perceived as having explicitly communicated its intention to promote diversity, nor as putting in place programs to develop it, or nor as developing positive attitudes and behaviors towards diverse people. The Type II refers to organizations that are perceived as displaying their intent to value diversity (such as signing a diversity commitment) without being perceived as implementing supporting programs nor changing attitudes and behaviors. Type III organizations may be perceived as having set up processes to promote diversity, but these are not perceived as being reflected into its employees’ attitudes and behaviors. This may characterize organizations that are concerned about complying with the law or other standards, but with no perceived success in doing so. It could also be an intermediary stage towards achieving a perceived change in behaviors and attitudes at some later stage. Type IV may characterize organizations that are perceived to ‘do diversity’, even without perceived explicit procedures or formal expression of intent. This could be considered as the final step of integrating diversity in some maturity life cycle approach of diversity management, where one, several, or all of configurations V to VIII may figure as intermediary stages. These organizational configurations are of course speculative, but they highlight the relevance of our study for future DC research and practice. The three-faceted DC conceptualization and operationalization along with the abovementioned configurations, emphasize the need for considering diversity management as a process rather than some fixed state (Carstens and De Kock, 2016). Diversity management should rather be seen as a succession of stages. Our work also raises several substantive questions. Are some organizations ‘born’ diversity friendly with no necessity to even intend to become so or without explicit supportive programs towards that end? How long can organizations durably remain in configurations II, III and V without generating tensions?

# Research limitations and future research

Our research has limitations that should be acknowledged. Common method bias is a potential issue that may have affected our findings. We encourage scholars to use multiple data collection methods in the future. Moreover, studies across more organizations would benefit from greater variability in employee perceptions and allow to better understand the role played by each of the three facets of DC on a range of outcomes. We also call for replication of our study within and/or organizations across several national contexts.

Despite these limitations, our study identifies a number of research avenues. We empirically tested the reconceptualization of DC proposed by Cachat-Rosset et al. (2017) and provided evidence of its robustness. Most of all, this research calls for further investigation of the concept. We invite future researchers to conduct further empirical tests of the newly-designed DC concept and use it in attempts to predict important outcomes such as turnover intentions, application intentions, job satisfaction or innovation as suggested by Cox (1993). This would also help to further assess the criterion validity of the measurement instrument. Nomological validity could also be further investigated by looking at other notions that are conceptually related to DC, including other organizational climates (such as innovation or ethical climates). More generally we encourage scholars to investigate the role of each facet of DC in influencing both individual and organizational outcomes. More practically, our three-dimensional measure of DC suggests that there are eight possible configurations for diversity management, drawing a preliminary tool-kit for practitioners in order to set up resistance-strategies against diversity status quo within their organization. What had previously been identified as distinct paradigms or perspectives (Cox, 1991; Dwertmann et al., 2016; Thomas, 1990; Thomas and Ely, 1996) may to some extent coincide with one or several of these configurations. Yet these configurations, the factors that cause firms to move between them, and their outcomes, remain to be studied on large samples of firms. This is without a doubt an ambitious research agenda.

# References

Agocs, C. (2002). Systemic discrimination in employment: mapping the issue and policy responses. In K. L. International (Ed.), *Workplace equality: international perspectives on legislation, policy and practice*. The Hague.

Anderson, J. C., and 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.

Arthur, J. B., & Boyles, T. (2007). Validating the human resource system structure: A levels-based strategic HRM approach. *Human Resource Management Review*, *17*(1), 77-92.

Bauer, T., and Green, S. (1998). Testing the combined effects of newcomer information seeking and manager behavior on socialization. *Journal of Applied Psychology*, *83*(1), 72–83.

Beus, J. M., Payne, S. C., Arthur, W., and Muñoz, G. J. (2017). The Development and Validation of a Cross-Industry Safety Climate Measure: Resolving Conceptual and Operational Issues. *Journal of Management*.

Bowen, D. E., & Ostroff, C. (2004). Understanding the HRM-firm performance linkages: The role of the “strength” of the HRM system. *Academy of Management Review*, 29(2), 203−221.

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

Buttner, E. H., Lowe, K. B., & Billings-Harris, L. (2012). An Empirical Test of Diversity Climate Dimensionality and Relative Effects on Employee of Color Outcomes. *Journal of Business Ethics*, 110(3), 247–258.

Cachat‐Rosset, G., Carillo, K., and Klarsfeld, A. (2017). Reconstructing the Concept of Diversity Climate – A Critical Review of Its Definition, Dimensions, and Operationalization. *European Management Review*.

Carstens, J. G., and De Kock, F. S. (2016). Firm-level diversity management competencies: development and initial validation of a measure. *The International Journal of Human Resource Management*, 1–27.

Chan, D. (1998). Functional relations among constructs in the same content domain at different levels of analysis: A typology of composition models. *Journal of Applied Psychology*, *83*(2), 234–246.

Chrobot-Mason, D., (2003). Keeping the promise. *Journal of Managerial Psychology,* 18 (1), 22 - 45

Chrobot-Mason, D., and Aramovich, N. P. (2013). The Psychological Benefits of Creating an Affirming Climate for Workplace Diversity. *Group and Organization Management*, *38*(6), 659–689.

Chung, B. G., Ehrhart, K. H., Shore, L. M., Randel, A. E., Dean, M. A., & Kedharnath, U. (2019). Work Group Inclusion: Test of a Scale and Model. *Group & Organization Management*. 1-28

Chung, Y., Liao, H., Jackson, S. E., Subramony, M., Colakoglu, S., & Jiang, Y. (2015). Cracking but Not Breaking: Joint Effects of Faultline Strength and Diversity Climate on Loyal Behavior. *Academy of Management Journal*, *58*(5), 1495–1515.

Churchill Jr., G. A. (1979). A Paradigm for Developing Better Measures of Marketing Constructs. *Journal of Marketing Research*, *16*(Feb), 64–73.

Commission européenne, et Direction générale de l’emploi, des relations industrielles et des affaires sociales Unité D.3. (2003). *Coûts et avantages de la diversité*. Luxembourg : Office des publications officielles des Communautés européennes.

Cox, T. (1993). *Cultural Diversity in Organizations: Theory, Research and Practice* (New edition). San Francisco, Calif.: Berrett-Koehler.

Cox, T. H., and Blake, S. (1991). Managing Cultural Diversity: Implications for Organizational Competitiveness. *The Executive*, *5*(3), 45–56.

Cox, T. (1991). The multicultural organization. *Executive*, *5*(2), 34–47.

Dwertmann, D. J., Nishii, L. H., and Van Knippenberg, D. (2016). Disentangling the Fairness and Discrimination and Synergy Perspectives on Diversity Climate: Moving the Field Forward. *Journal of Management*, *42*(5), 1136–1168.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. (Sage Publications Ltd., Ed.).

Fiske, S. T., & Taylor, S. E. (1991). Social cognition. *Mcgraw-Hill Book Company*.

Fornell, C., and Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 39–50.

Glick, W. H. (1985). Conceptualizing and Measuring Organizational and Psychological Climate: Pitfalls in Multilevel Research. *Academy of Management Review*, *10*(3), 601–616.

Gonzalez, J. A. and A. S. Denisi (2009). “Cross-level effects of demography and diversity climate on organizational attachment and firm effectiveness”. *Journal of Organizational Behavior*, 30: 21–40.

Guchait, P., J. Madera and M. Dawson (2016). “Learning in the service environment: the influence of diversity climate”. *Journal of Service Theory and Practice*, 26(4): 448–470.

Guzzo, R. A., Yost, P. R., Campbell, R. J., and Shea, G. P. (1993). Potency in groups: Articulating a construct. *British Journal of Social Psychology*, *32*(1), 87–106.

Hair, J. F., Hult, G. T. M., Ringle, C. M., and Sarstedt, M. (2017). 2nd Ed., Sage: Thousand Oaks.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a Silver Bullet. *The Journal of Marketing Theory and Practice*, *19*(2), 139–152.

Herring, C. (2009). Does diversity pay? Race, gender, and the business case for diversity. *American Sociological Review*, *74*, 208–224.

Hicks-Clarke, D. and P. Iles (2000). “Climate for diversity and its effects on career and organisational attitudes and perceptions”. *Personnel Review*, 29: 324–345.

Hobman, E. V., Bordia, P., & Gallois, C. (2004). Perceived dissimilarity and work group involvement: The moderating effects of group openness to diversity. *Group & Organization Management*, 29(5), 560-587.

Hooper, D., Coughlan, J., & Mullen, M. (2008). *Structural equation modelling: Guidelines for determining model fit*. Articles, 2.

Hopkins, W. E., Hopkins, S. A., and Mallette, P. (2001). Diversity and managerial value commitment: A test of some proposed relationships. *Journal of Managerial Issues*, *13*(3), 288–306.

Hox, J. (2010). *Multilevel analysis: Techniques and applications*. (Routledge, Ed.) (2nd ed.). New York, NY.

Katz, D., and Kahn, R. L. (1978). *The Social Psychology of Organizations*. (Vol. 2, p. 528). New York: Wiley.

Klein, K. J., and Kozlowski, S. W. (2000). *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions.* (Jossey-Bas, Ed.).

Kline, P. (2013). *Handbook of psychological testing.* (Routledge., Ed.).

Kopelman, R. E., A. P. Brief and R. A. Guzzo, 1990, “The role of climate and culture in productivity”. *Organizational Climate and Culture*, 282–318.

Kossek, E. E., and Zonia, S. C. (1993). Assessing diversity climate: A field study of reactions to employer efforts to promote diversity. *Journal of Organizational Behavior*, *14*(1), 61–81.

Kossek, E. E., Markel, K. S., and McHugh, P. P. (2003). Increasing diversity as an HRM change strategy. *Journal of Organizational Change Management*, *16*(3), 328–352.

Kulik, C. T. (2014). Working below and above the line: the research–practice gap in diversity management. *Human Resource Management Journal*, 24(2), 129-144.

Law, K. S., Wong, C. S., and Mobley, W. M. (1998). Toward a taxonomy of multidimensional construct. *Academy of Management Review*, *23*(4), 741–755.

LeBreton, J. M., and Senter, J. L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods*, *11*(4), 815–852.

Mackenzie, S. B., Podsakoff, P. M., and Podsakoff, N. P. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS Quarterly*, *35*(2), 293–334.

Madera, J. M., M. Dawson and J. A. Neal (2013). “Hotel managers’ perceived diversity climate and job satisfaction: The mediating effects of role ambiguity and conflict”. *International Journal of Hospitality Management*, 35: 28–34.

McKay, P. F., and Avery, D. R. (2015). Diversity climate in Organizations: Current Wisdom and Domains of Uncertainty. *Research in Personnel and Human Resources*, *33*, 191–233.

McKay, P. F., Avery, D. R., & Morris, M. A. (2009). A Tale of Two Climates: Diversity Climate from Subordinates’ and Managers’ Perspectives and their Role in Store Unit Sales Performance. *Personnel Psychology*, 62(4), 767-791.

McKay, P. F., Avery, D. R., and Morris, M. A. (2008). Mean racial‐ethnic differences in employee sales performance: The moderating role of diversity climate. *Personnel Psychology*, *61*(2), 349–374.

Mischel, W. (1973). Toward a cognitive social learning reconceptualization of personality. *Psychological review*, *80*(4), 252.

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 and Governance*, *40*(4), 305–333.

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

Mor Barak, M. E., Cherin, D. A., & Berkman, S. (1998). Organizational and Personal Dimensions in Diversity Climate Ethnic and Gender Differences in Employee Perceptions. *The Journal of Applied Behavioral Science*, *34*(1), 82–104.

Moore, G., and Benbasat, I. (1991). Development of an Instrument to Measure the Perceptions of Adopting an Information Technology Innovation. *Information Systems Research*.

Morrison, E. W. (1993). Longitudinal study of the effects of information seeking on newcomer socialization. *Journal of Applied Psychology*, *78*(2), 173–183.

Nishii, L. H. (2013). The benefits of climate for inclusion for gender-diverse groups. *Academy of Management Journal*, 56(6), 1754-1774.

Nishii, L. H., Lepak, D. P., & Schneider, B. (2008). Employee attributions of the “why” of HR practices: Their effects on employee attitudes and behaviors, and customer satisfaction. *Personnel Psychology*, *61*(3), 503-545.

Nunnally, J.C. (1978). *Psychometric Theory* (2nd ed.). New York: McGraw-Hill.

Ozgen, C., Nijkamp, P., and Poot, J. (2013). The impact of cultural diversity on firm innovation: evidence from Dutch micro-data. *IZA Journal of Migration*, *2*(1), 18.

Price, J. L., and Mueller, C. W. (1986). Absenteeism and turnover of hospital employees.

Reichers, A., and Schneider, B. (1990). *Climate and culture: An evolution of constructs*. *Cunha, MP et al., op. cit.* San Francisco.

Reinwald, M., Huettermann, H., & Bruch, H. (2018). Beyond the mean: Understanding firm‐level consequences of variability in diversity climate perceptions. *Journal of Organizational Behavior*. 2019; 1– 20.

Richard, O. C., Roh, H., and Pieper, J. R. (2013). The Link Between Diversity and Equality Management Practice Bundles and Racial Diversity in the Managerial Ranks: Does Firm Size Matter? *Human Resource Management*, *52*(2), 215–242.

Roberson, Q. M. (2006). Disentangling the Meanings of Diversity and Inclusion in Organizations. *Group and Organization Management*, *31*(2), 212–236.

Schmiedel, T., Vom Brocke, J., and Recker, J. (2014). Development and validation of an instrument to measure organizational cultures’ support of business process management. *Information and Management*, *51*(1), 43–56.

Schneider, B., 1990. Organizational climate and culture. San Francisco: Pfeiffer.

Schneider, B., and Reichers, A. E. (1983). On the etiology of climates. *Personnel Psychology*, *36*(1), 19–39.

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

Spence, M. (1974). “Competitive and optimal responses to signals: An analysis of efficiency and distribution”. *Journal of Economic Theory*, 7: 296–332.

Straub, D. W. (1989). Validating instruments in MIS research. *MIS Quarterly*, 147–169.

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. G. Austin (Eds.), *Psychology of intergroup relations* (2nd ed., pp. 7-24). Chicago, IL: Nelson-Hall.

Thomas Jr., R. R. (1990). From Affirmative Action to Affirming Diversity. *Harvard Business Review*, *68*(2), 107–117.

Thomas, D., and Ely, R. (1996). Making differences matter: a new paradigm for diversity management. *Harvard Business Review, Vol. 74, N° 5, p. 79-90.*, *74*(5), 79–90.

Van Campenhoudt, L. and R. Quivy (2011). Manuel de recherche en sciences sociales. Paris : Dunod.

Van der Vegt, G. S., and Janssen, O. (2003). Joint Impact of Interdependence and Group Diversity on Innovation. *Journal of Management*, *29*(5), 729–751.

Vertovec, S. (2007). Super-diversity and its implications. *Ethnic and Racial Studies*, *30*(6), 1024–1054.

Wanberg, C. R., and Kammeyer-Mueller, J. D. (2000). Predictors and Outcomes of Proactivity in the Socialization Process. *Journal of Applied Psychology*, *1*(3), 373–38.

# Tables

Table I: Grand Diversity management theoretical frameworks

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Insistence on…** | **Cox (1991)** | **Thomas (1990)** | **Thomas & Ely (1996)** | **Dwertmann et al. (2016)** |
| **…fair representation and treatment** | Plural type | Affirmative Action | Discrimination & Fairness | Fairness & Discrimination |
| **…valuing diversity** | Multi-cultural type | Affirming Diversity | Access & Legitimacy  Learning & Effectiveness | Synergy |

Table II: Diversity climate conceptualization

|  |  |  |
| --- | --- | --- |
| **Diversity climate** | | |
| *the shared perception by employees of intent, programs, attitudes and behaviors in favor of relevant forms of diversity in the social context of the organization* | | |
| **Intentionality** | **Programmatic** | **Praxis** |
| *perceptions that the organization values diversity in discourse* | *perceptions of formal diversity programs or policies and action plans set up within the organization to develop and support workforce diversity* | *perceptions about management’s and colleagues’ attitudes and behaviors towards diverse people* |

Table III: Limitations inherent in diversity climate measures

|  |  |
| --- | --- |
| **Limitations** | **Review where limitation is initially identified** |
| Dominance of the fairness & discrimination over the synergy perspective | Dwertmann et al. (2016) |
| Misalignment between concept definition and measure dimensions/items | Chrobot-Mason and Aramovich (2013) |
| Individual level analysis while relying on organizational climate framework | McKay and Avery (2015) |
| Target of perceptions varied (commitment, policies, behaviors such as valuing, listening, etc…) | Dwertmann et al. (2016) |
| Use of demographic data rather than perceptions | Cachat-Rosset et al. (2017) |
| Foci not relevant with organizational climate assessment (self, designated categories).  Designated categories hamper cross-national comparisons. | Cachat-Rosset et al. (2017) |

Table IV: Stages presentation

|  |  |  |  |
| --- | --- | --- | --- |
| **Stage / Step** | **Method / Technique** | **Number of resulting items** | **Actor/**  **Sample** |
| **Stage 1: Scale Development** | | |  |
| Construct dimensionality | Conceptual development |  | |
| Items generation | Literature review | 114 | 34 articles |
| Item selection | Card sorting | 26 | 3 judges |
| Item purification | Expert panel | 26 | 4 experts |
| **Stage 2: Scale validation at individual level** | | | |
| Further item purification & Initial validity assessment | EFA  internal consistency  scale reliability  indicator reliability convergent validity discriminant validity | 12 | Study 1  N = 150 |
| Validity assessment | CFA  internal consistency  scale reliability  indicator reliability convergent validity discriminant validity Model fit | 11 | Study 2  N = 398 |
| **Stage 3: Scale validation at aggregate level** | | | |
| Nomological assessment | Hierarchical Linear Modelling (HLM) | 11 | Study 3  N = 221 |
| Model fit at aggregate level | CFA | 11 |
| Predictive validity | Hierarchical Linear Modelling (HLM) using McKay et al. (2008) measure | 11 |

Table V: inter-judge agreement on items (closed card-sorting round)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Intentionality** | **Programmatic** | **Praxis** | **Not adapted or unclear** | **Total** | **% Fit** |
| **Intentionality** | 30 | 0 | 2 | 2 | 34 | **88%** |
| **Programmatic** | 1 | 49 | 4 | 0 | 54 | **91%** |
| **Praxis** | 6 | 11 | 55 | 2 | 74 | **74%** |
| **Not adapted or unclear** | 3 | 10 | 15 | 38 | 66 | **58%** |
|  | **Total fit** | **172** |  | **Total placement** | **228** | **75,4%** |

Table VI: Reliability and Average Variance Extracted

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | CR | C’s α | Rho | AVE |
| INT | 0.860 | 0.782 | 0.795 | 0.607 |
| PRAX | 0.846 | 0.757 | 0.760 | 0.579 |
| PROG | 0.892 | 0.839 | 0.842 | 0.674 |

Table VII: Item loadings

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Item Code** | **Item** | **INT** | **PROG** | **PRAX** |
| INT1 | Top management is visibly committed to diversity | **0.805\*\*\*** | 0.462 | 0.432 |
| INT2 | Top management intends to have a diverse workforce | **0.860\*\*\*** | 0.333 | 0.446 |
| INT3 | Clearly, diversity is not important to this organization | **0.650\*\*\*** | 0.388 | 0.239 |
| INT4 | This organization values differences among employees | **0.781\*\*\*** | 0.332 | 0.300 |
| PROG1 | This organization has set up action plans to promote diversity | 0.383 | **0.778\*\*\*** | 0.147 |
| PROG2 | This organization has launched initiatives to diversify its workforce | 0.408 | **0.866\*\*\*** | 0.131 |
| PROG3 | There are programs targeting underrepresented groups | 0.319 | **0.707\*\*\*** | 0.080 |
| PROG4 | There are policies promoting diversity in the workplace | 0.421 | **0.873\*\*\*** | 0.092 |
| PRAX1 | Discrimination is not tolerated in this organization | 0.329 | 0.095 | **0.636\*\*\*** |
| PRAX2 | Discriminatory behaviors are discouraged in this organization | 0.360 | 0.166 | **0.667\*\*\*** |
| PRAX3 | Managers work well with all kind of employees without distinction | 0.348 | -0.042 | **0.796\*\*\*** |
| PRAX4 | Workgroups welcome employees who are different | 0.374 | 0.153 | **0.869\*\*\*** |
| *Significance level: \*\*\* p < 0.001 / \*\* p < 0.01 / \* p< 0.05* | | | | |

Table VIII: Correlations and square roots of AVE on the diagonal and HTMT ratios

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Correlations and square roots of AVE on the diagonal | | | HTMT ratios | |
| INT | PRAX | PROG | INT | PRAX |
| INT | 0.779 |  |  |  |  |
| PRAX | 0.464 | 0.760 |  | 0.588 |  |
| PROG | 0.478 | 0.133 | 0.821 | 0.589 | 0.199 |

Table IX: Reliability and Average Variance Extracted

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | CR | C’s α | Rho | AVE |
| INT | 0.887 | 0.792 | 0.889 | 0.707 |
| PRAX | 0.845 | 0.755 | 0.760 | 0.579 |
| PROG | 0.929 | 0.897 | 0.905 | 0.765 |

Table X: Item loadings

|  |  |  |  |
| --- | --- | --- | --- |
| **Item code** | **INT** | **PROG** | **PRAX** |
| INT1 | **0.872\*\*\*** | 0.538 | 0.551 |
| INT2 | **0.824\*\*\*** | 0.548 | 0.415 |
| INT3 | Dropped | | |
| INT4 | **0.826\*\*\*** | 0.432 | 0.570 |
| PROG1 | 0.544 | **0.913\*\*\*** | 0.329 |
| PROG2 | 0.569 | **0.902\*\*\*** | 0.355 |
| PROG3 | 0.381 | **0.818\*\*\*** | 0.262 |
| PROG4 | 0.593 | **0.865\*\*\*** | 0.376 |
| PRAX1 | 0.451 | 0.286 | **0.827\*\*\*** |
| PRAX2 | 0.362 | 0.331 | **0.725\*\*\*** |
| PRAX3 | 0.426 | 0.230 | **0.755\*\*\*** |
| PRAX4 | 0.595 | 0.307 | **0.734\*\*\*** |

Table XI: Correlations and square roots of AVE on the diagonal and HTMT ratios

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Correlations and square roots of AVE on the diagonal | | | HTMT ratios | |
| INT | PRAX | PROG | INT | PRAX |
| INT | 0.841 |  |  |  |  |
| PRAX | 0.610 | 0.761 |  | 0.778 |  |
| PROG | 0.602 | 0.381 | 0.875 | 0.706 | 0.458 |

Table XII: Descriptive statistics and correlations among variables

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Variables | Mean | SD | 1 | 2 | 3 | 4 |
| Gender | 1.56 | 0.53 |  |  |  |  |
| Management | 1.70 | 0.46 | 0.146\* |  |  |  |
| Social integration | 4.16 | 0.65 | -0.169\* | -0.022 |  |  |
| Diversity climate  (new measure) | 3.53 | 0.15 | -0.184\*\* | -0.023 | 0.144\* |  |
| DC (McKay et al., 2008) | 3.86 | 0.16 | -0.308\*\* | -0.184\* | 0.115\* | 0.896\*\*\* |

*N = 11 at unit level; N = 221 at individual level – \*\*\* p < 0.001 \*\* p < 0.01 \* p < 0.05*

Table XIII: Results of HLM analysis on social integration

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Variable (coefficient)** | **Model 1 (Null)** | | | **Model 2 (Fixed)** | | | **Model 3**  **(measure from McKay et al., 2008)** | | |
|  | b | SE | t | b | SE | t | b | SE | t |
| Intercept (SI)  Unit diversity climate | 4.161 | 0.044 | 94.35\*\*\* | 4.162  0.570 | 0.039  0.075 | 105.9\*\*\*  7.60\*\*\* | 4.28  0.422 | 0.044  0.069 | 97.8\*\*\*  6.10\*\*\* |
| Deviance (-2 log likelihood) | 433.9 | | | 395.7 | | | 170.2 | | |
| Δ(-2 log likelihood) |  | | | 38.2\*\*\* | | |  | | |
| AIC | 447.9 | | | 399.8 | | | 174.4 | | |
| BIC | 454.7 | | | 406.5 | | | 179.8 | | |
| R² | 0.055 | | | 0.287 | | | 0.230 | | |
| Δ R² |  | | | 0.232 | | |  | | |

*N = 11 at unit level; N = 221 at individual level – \*\*\* p < 0.001 \*\* p < 0.01 \* p < 0.05*

Table XIV: Unit-level model fit assessment

|  |  |  |
| --- | --- | --- |
| **Model / Fit statistics** | **Model A 🡪 SI** | **Model B 🡪 SI** |
| χ²  CMIN/df  CFI  SRMR  RMSEA | 196.5\*\*\*  2.29  0.90  0.072  0.076 | 162.1\*\*\*  1.977  0.93  0.068  0.067 |

Table XV: Diversity climate taxonomy

|  |  |  |  |
| --- | --- | --- | --- |
| **Organizational configuration** | **Intentionality** | **Programmatic** | **Praxis** |
| I | Low | Low | Low |
| II | **High** | Low | Low |
| III | Low | **High** | Low |
| IV | Low | Low | **High** |
| V | **High** | **High** | Low |
| VI | Low | **High** | **High** |
| VII | **High** | Low | **High** |
| VIII | **High** | **High** | **High** |

Figure 1: Model Fit Assessment (N= 398)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **MODEL A: DC as a second-order reflexive construct** | | | | | | | |
|  | | | | | | | |
| *χ²= 167.0*  *p < 0.000* | *CMIN/df = 4.07* | *CFI = 0.943* | | *SRMR = 0.072* | | *RMSEA = 0.088* |
| **MODEL B: DC as a first-order reflexive construct** | | | | | | | |
|  | | | | | | | |
| *χ² = 525.9*  *p < 0.000* | *CMIN/df = 12.52* | *CFI = 0.782* | *SRMR = 0.120* | | *RMSEA = 0.170* | |