Running head. Multilevel Job Demands and Resources

A Multilevel Perspective on the Role of Job Demands, Job Resources, and Need Satisfaction for Employees' Outcomes

Alexandre J.S. Morin*
Substantive-Methodological Synergy Research Laboratory, Concordia University
Nicolas Gillet*
Université de Tours, QualiPsy EE 1901, Tours, France
& Institut Universitaire de France (IUF)
Ann-Renée Blais*
Statistics Canada
& Department of National Defence, Canada
Caitlin Comeau
Department of National Defence
Simon A. Houle
Substantive-Methodological Synergy Research Laboratory, Concordia University

* The first three authors (A.J.S.M., N.G., & A.-R.B.) contributed equally to this article and their order was determined at random: All three should be considered first authors.

This is the prepublication version of the following manuscript:
Morin, A.J.S., Gillet, N., Blais, A.-R., Comeau, C., & Houle, S.A. (2023, In Press). A multilevel perspective on the role of job demands, job resources, and need satisfaction employees' outcomes. *Journal of Vocational Behavior*. https://doi.org/10.1016/j.jvb.2023.103846.

Acknowledgements. Preparation of this paper was supported by grants from the Canadian Institute for Military and Veteran Health Research (CIMVHR) and from the Social Sciences and Humanities Research Council of Canada (435-2018-0368).

Corresponding author:
Alexandre J.S. Morin, Substantive-Methodological Synergy Research Laboratory, Psychology, Concordia University, 7141 Sherbrooke W, Montreal, QC, Canada, H3B 1R6
Email: alexandre.morin@concordia.ca

Abstract

This study investigates the mediator role of psychological need satisfaction for the effects of job demands and resources on turnover intentions, psychological distress, and work-to-family conflict, simultaneously at the employee and work unit levels. In doing so, we consider how need satisfaction, when considered at the work unit level, creates a context likely to play an additional role in the prediction of these outcomes. These questions were investigated using a combination of doubly latent multilevel confirmatory factor analyses and structural equation models applied to responses provided by a large sample (N = 5,716 employees nested within 50 work units) of Canadian Armed Forces/Department of National Defence personnel. The results supported the idea that work environment effects on the outcomes considered in this study were mediated by psychological need satisfaction at the individual and work unit levels and demonstrated that these associations were driven by global work environment perceptions and global need satisfaction. Furthermore, need satisfaction was found to create a context, at the work unit level, leading employees working in units including more highly satisfied co-workers to present higher levels of turnover intentions but lower levels of work-to-family conflict than would be expected based on their individual levels of need satisfaction.

Keywords: Doubly Latent; Multilevel Measurement; Latent Aggregation; Context; Climate; Centering; Multilevel Mediation; Job Demands and Resources; Need Satisfaction; Psychological Distress; Turnover Intentions; Work-to-Family Conflict.

Highlights

- The effects of job demands, job resources, and need satisfaction are considered across levels (individual *vs* group)
- Data came from a large sample of (N = 5,716 employees, 50 work units) of Canadian Armed Forces personnel
- Work environment effects on employees were mediated by psychological need satisfaction across levels
- In more satisfied units, employees presented higher levels of turnover intentions but lower levels of work-to-family conflict

The Job Demands-Resources (JD-R) model (Demerouti et al., 2001) is arguably one of the major framework for studying the influence of work characteristics on employees (Bakker & Demerouti, 2017; Schaufeli & Taris, 2014). This model highlights the role of two categories of work conditions, job demands and resources. Job demands refer to aspects of a job that require sustained physical and/or psychological effort and are associated with physiological and/or psychological costs. In contrast, job resources refer to job characteristics that support employees in achieving their goals, help reduce the costs of job demands, and stimulate development. Although research has generally supported the basic assumptions of this model (e.g., Bakker & Demerouti, 2007, 2017; Podsakoff et al., 2007), JD-R research presents some critical limitations that the present study seeks to address.

First, Bakker and Demerouti (2017) recently reinforced the importance for JD-R research to systematically adopt a multilevel perspective to obtain a more accurate, and practically useful, picture of the role played by shared perceptions of job demands and resources, reflecting the common work environment to which all members of a work unit, workgroup, or organization are jointly exposed, relative to that played by inter-individual differences in these perceptions. In this study, we rely on doubly latent multilevel confirmatory factor analyses (DL-MLCFA) and structural equation models (DL-MLSEM; Morin et al., 2022) to differentiate these two layers of influence underlying the role of job demands (role ambiguity and work overload) and of job resources linked to the leader (contingent recognition and reward, transformational leadership, and safety actions and expectations), workgroup (psychological safety and interpersonal justice), and organization (support and culture of interpersonal respect) for employees' turnover intentions, psychological distress, and work-to-family conflict. This complementary set of differentiated outcomes were selected because they are known to carry important costs in terms of individual, work unit, and organizational performance (e.g., Hancock et al., 2013; Lim & Tai, 2014; Odle-Dusseau et al., 2012), in addition to providing a broad overview of the impact of job demands and resources on employees' functioning at work (turnover intentions), in general (psychological distress), and in their familial life (work-to-family conflict). In particular, work-to-family conflict occurs when job demands interfere with one's ability to meet family demands (Odle-Dusseau et al., 2012), and thus come to take a toll on employees' personal life (Zhou et al., 2019). The JD-R model (e.g., Bakker & Demerouti, 2017; Demerouti et al., 2001; Schaufeli, & Taris, 2014) leads us to expect a similar impact of work characteristics on these three outcomes, although this impact is likely to be stronger for characteristics more directly related to one's work life (turnover intentions) than for domain-general characteristics (psychological distress) or characteristics reflecting the interference of work on one's personal life (work-to-family conflict).

Second, although the JD-R model provides a comprehensive heuristic for the study of work characteristics, it falls short of identifying the psychological mechanism(s) explaining their effects on employees (Schaufeli & Taris, 2014). This limitation has led to the reliance on other theories to study these mechanisms (Bakker & Demerouti, 2017; Schaufeli & Taris, 2014). In this study, we rely on Self-Determination Theory (SDT; Ryan & Deci, 2017) to consider the role of psychological need satisfaction as a mediator (Van den Broeck et al., 2008) of the associations between job demands and resources and our outcomes as they occur at the individual and work unit levels. In doing so, we propose and test a workgroup-level theorization whereby need satisfaction can foster a distinctive work context at the work unit level, in addition to its documented individual benefits.

Third, the bulk of research on job demands and resources has considered them either from a one-dimensional perspective (e.g., Boudrias et al., 2011, 2014), thus ignoring the possibility that distinct demands and resources can play a unique role in prediction, or from a multidimensional perspective (e.g., Boudrias et al., 2020; Kubicek et al., 2015), thus ignoring their common core. This second approach makes it difficult to assess the unique role of distinct work characteristics because of their moderately high inter-correlations (e.g., Bosmans et al., 2019; De Gieter et al., 2018). In the present study, we start by conducting a complete assessment of the multilevel measurement structure of work characteristics to achieve a proper estimate of the role played by their common core (global work environment quality) and of the unique role played by each work characteristic beyond this commonality. We then build upon this optimal measurement structure to consider the multilevel associations between the work environment characteristics and outcomes considered in this study, as well as the mediator role of employees' own levels of need satisfaction (individual level, L1) and of the need satisfaction context formed by the extent to which needs are satisfied among members of a work unit (work unit level, L2).

A Multilevel Multidimensional Perspective on the Role of Job Demands and Resources

An Inherently Multilevel Perspective

The theoretical propositions of the JD-R model (Demerouti et al., 2001) have received a fair amount of empirical support, confirming the negative role of job demands and the positive role of job resources for a variety of outcomes related to employees' health, attitudes, and behaviors, including the outcomes of the present study (i.e., turnover intentions, psychological distress, and work-to-family conflict; e.g., Bakker & Demerouti, 2007, 2017; Podsakoff et al., 2007). However, although substantial JD-R research has focused on individual resources (e.g., self-efficacy or optimism; e.g., Xanthopoulou et al., 2007), research suggests that job resources may play an even greater role for employees' health, attitudes, and behaviors (e.g., Bosmans et al., 2019).

Although the JD-R model has been multilevel from its inception (Bakker & Demerouti, 2017; Demerouti et al., 2001), the bulk JD-R research has primarily focused on individual perceptions of job characteristics, ignoring the fact that members of a work unit or an organization may share similar perceptions reflecting the influence of common external features of their work environment. It is well established in the organizational sciences (Gonzàlez-Romà & Hernàndez, 2017) that conclusions obtained at one level of analysis (e.g., based on employees' perceptions) cannot be expected to generalize at other levels of analysis (e.g., the work unit). To better capture the multilevel nature of job resources, Nielsen et al. (2017) proposed to classify them according to their origin (i.e., leader, workgroup, or organization) and showed that each type of resource played a complementary, and yet systematically positive, role in relation to employees' health, attitudes, and behaviors. Workgroup resources relate to the horizontal social and interpersonal context of the workplace; that is, to relationships among workgroup members that can help foster an efficient exchange of information, positive social interactions, and feelings of trust and fairness. Leader resources refer to vertical interactions between employees and their supervisors who may, by virtue of their position and leadership, provide support, guidance, and security. Finally, organizational resources refer to the broader work context, and the way it is organized and managed to support and motivate employees' growth and functioning.

A Broad, Heuristic, and Multidimensional Perspective

Nielsen et al.'s (2017) contribution highlights the unrealism of trying to characterize the complex and multifactorial nature of the work environment via a restricted focus on a subset of characteristics, an idea that was previously expressed by Grawitch and Ballard (2016) when trying to capture the defining characteristics of psychologically healthy workplaces. These observations also highlight one key strength of the JD-R model, which provides a heuristic allowing researchers and practitioners to anticipate the likely impact of any job characteristics based on its classification as a demand or resource. Conversely, this strength is also a limitation if we consider that most JD-R research has considered only a small subset of job demands and resources. First, some research has simply relied on global measures of job demands and resources (e.g., Boudrias et al., 2011, 2014), thus completely ignoring the possible unique role of specific types of work characteristics beyond what they share with one another. Second, some research has specifically focused on distinct work characteristics (e.g., Kubicek et al., 2015), without considering their common core. Beyond these two extremes, research has also focused on the role played by a limited number (one or two) of work characteristics, making it impossible to obtain a global picture of the overarching work environment (e.g., Boudrias et al., 2020; Chen et al., 2020; Griffin, 2010). When we consider correlations among employees' perceptions of work characteristics, we know that distinct characteristics present both a common core as well as unique components, as highlighted by their moderately high inter-correlations (Bosmans et al., 2019; De Gieter et al., 2018). Importantly, in multilevel analyses, even when employees hold reasonably distinct perceptions of different constructs, this distinction does not necessarily translate to their aggregated perceptions, which sometimes come to reflect the global functioning of their workgroup or organization across all measures (i.e., a global factor encompassing all measures), rather than their distinctive nature (Morin et al., 2014, 2022).

These observations raise important questions regarding: (a) whether measures of job characteristics retain specificity beyond the assessment of employees' perceptions of their global work environment; and (b) whether employees' perceptions of their work environment form one overarching dimension (i.e., global levels of work environment quality) while retaining specificities uniquely associated with each job characteristic, or whether these job characteristics simply form correlated dimensions without

a common core (e.g., Morin et al., 2017). These two possibilities suggest that employees could perceive their work environment in a global manner, while also being able to highlight how their exposure to specific job characteristics deviate from this global perception. In this study, we rely on a bifactor representation of work environment perceptions to achieve a proper disaggregation of participants' global and specific perceptions of their work environment (Morin et al., 2017, 2022).

A Perspective Anchored in Practical Relevance

Any investigation of the JD-R model faces the critical challenge of selecting a comprehensive, and yet limitative, set of job demands and resources in a way that is, if not necessarily theoretically-driven, at least theoretically-coherent with the scope of the model. Lacking any clear guideline from the JD-R model regarding which, among many, work characteristics might be most relevant to consider (Demerouti et al., 2001; Nielsen et al., 2017), we rely on a theory-oriented abductive approach (Antonakis, 2017; Martela, 2015) anchored in a big data perspective (Toninandel et al., 2016, 2018). More precisely, our approach is "loosely" guided by the JD-R model, which highlights the relevance of job demands and of job resources linked to the workgroup, leader, and organization (Demerouti et al., 2001; Nielsen et al., 2017), and seeks to guide the refinement of this theoretical perspective without being impeded by its current lack of precision (e.g., Antonakis, 2017; Martela, 2015). This flexible approach allows to us to capitalize on a large organizational data archive of military personnel (Ivey et al., 2018), which was not collected for the specific purposes of this investigation, and yet provides rich information likely to inform further JD-R developments (e.g., Toninandel et al., 2016, 2018). More precisely, we rely on the results from comprehensive organizational assessment efforts realized in the Canadian Armed Forces/Department of National Defence (CAF/DND). This organizational assessment (the Unit Morale Profile version 2.0; UMP2) was devised by CAF/DND to operationalize the new national standards for health and safety in the workplace (Canadian Standards Association Group and Bureau de normalisation du Québec [CSA/BNQ], 2013). What makes the UMP2 so interesting is that it was created to operationalize these standards from the perspective of the JD-R model (Demerouti et al., 2001), while capturing job demands and job resources related to the workgroup, supervisor, and organization (Nielsen et al., 2017) that were deemed to be most important for CAF/DND personnel (Ivey et al., 2018). As such, relying on the UMP2 allows us to rely on practical relevance to compensate for a lack of theoretical guidance related to the choice of the optimal set of work characteristics to consider in our analyses. The work characteristics covered in this study are presented, and defined, in Table 1.

The Mediator Role of Psychological Need Satisfaction from a Multilevel Perspective

Despite its interest, the JD-R model (Demerouti et al., 2001) fails to clearly identify the psychological mechanism(s) explaining the effects of job demands and resources on employees (Schaufeli & Taris, 2014). For this reason, JD-R research typically relies on distinct theories to study these mechanisms (Bakker & Demerouti, 2017; Schaufeli & Taris, 2014). Some of these theoretical enrichments have led to the recognition that part of the effects of job demands and resources are mediated by work engagement (Huynh et al. 2014), emotional exhaustion (Knudsen et al., 2009) or work recovery (Kinnunen et al., 2011). Yet, this piecemeal approach cannot bridge the theoretical gap related to the core psychological mechanisms underpinning the effects of job demands and resources.

SDT has also become a dominant framework for the investigation of the role of environmental characteristics on employees' health, attitudes, and behaviors across life domains (Ryan & Deci, 2017), including work (Deci et al., 2017; Van den Broeck et al., 2016). A key contribution of SDT lies in the identification of psychological mechanisms underpinning the effects of work characteristics (Bakker & Demerouti, 2017; De Gieter et al., 2018). More precisely, SDT positions the satisfaction of employees' basic needs for autonomy (i.e., to experience of sense of volition of ownership of one's actions), competence (i.e., to feel a sense of mastery underpinning one's actions), and relatedness (i.e., to achieve a sense of belongingness and connection) as a core mechanism underpinning the action of work characteristics (Ryan & Deci, 2017). Given this complementarity, formal connections between the JD-R model and SDT have been proposed (Van den Broeck et al., 2008), positioning need satisfaction as a mechanism underlying the action of job demands and resources. This joint JD-R/SDT framework has been supported in research focusing on psychological health (Boudrias et al., 2011, 2014), engagement and exhaustion (Gillet et al., 2015), and performance (De Gieter et al., 2018).

The mediator role of need satisfaction has rarely been investigated in a multilevel context, although emerging evidence suggests that this mediator role may hold across levels. For instance, in a study of

university students, Kelly et al. (2008) studied the group-level impact of need satisfaction. Their results showed that students from groups characterized by higher levels of need satisfaction reported better outcomes (satisfaction and positive affect) than students from groups characterized by lower levels of need satisfaction. Students whose need satisfaction was higher than their group's average also reported better outcomes than students characterized by levels of need satisfaction lower than their group average. In a study of employees, Schreurs et al. (2014) showed that teams characterized by more intrinsic, relative to extrinsic, work values demonstrated higher work engagement than teams characterized by more extrinsic, relative to intrinsic, work values. However, although these authors also showed that this relation was mediated by need satisfaction, the second part of this mediation (need satisfaction – work engagement) was only assessed at the individual level. Lastly, Jungert et al. (2018) showed that the benefits of a group-based need supportive intervention on motivation were mediated by need satisfaction. However, although the group-based nature of this intervention may suggests group-levels effects of need satisfaction, the second part of this mediation (need satisfaction – outcomes) was also only assessed at the individual level. Thus, despite their interest, none of those studies considered the whole JD-R/SDT framework as it occurs across levels.

Two Types of Multilevel Constructs

Two types of multilevel constructs can be identified when employees' ratings are used to assess constructs at the individual (L1) and workgroup (L2) levels (Gagné et al., 2020; Marsh et al., 2012; Morin et al., 2014, 2022). This distinction is related to the referent of the ratings. *Contextual* constructs are formed by the aggregation of individual ratings in which the referent is the person (L1). Thus, a contextual construct entails the assessment of a meaningful individual characteristic (e.g., gender) that takes on a different meaning at the workgroup-level (e.g., gender composition of the workgroup). In contrast, *Climate* constructs are formed by the aggregation of individual ratings in which the referent is the workgroup (L2)[1]. A climate constructs thus relies on the direct assessment of their workgroup reality by the employees, and the separation of these ratings into two components reflecting employees' shared perceptions of the workgroup reality (L2) and inter-individual differences in these perceptions (L1).

This distinction is only relevant when constructs are considered as predictors (e.g., Morin et al., 2022). Statistically, acknowledging that individual measures can be impacted by (be outcomes of) L1 and L2 predictors requires that they also be modelled at both levels (i.e., workgroup-level constructs can only predict outcome variance at the same level; Preacher et al., 2010). Workroup-level predictions reflect the shared impact of the workgroup reality across all members of the workgroup, without changing the meaning of the outcome being assessed. When relying on multilevel analyses, the analytic model automatically separates outcome variance across two levels: The L2 component reflects the shared influence of the predictor across members of the workgroup, whereas the L1 component captures how inter-individual differences in predictor levels predict inter-individual differences in outcome levels (Morin et al., 2014, 2022).

When studying the predictor role of job demands and resources, which are conceptualized in the JD-R model as characteristics of the work environment, employees are (or at least should be) asked to directly describe their work environment (Marsh et al., 2012; Morin et al., 2022). When considered as predictors, these ratings form *climate* constructs at L2, reflecting shared employees' perceptions about the reality of their work units. With *climate* constructs, the residual L1 component of these ratings reflects inter-individual differences in workgroup members' perceptions of the L2 reality. In plainer language, when the referent of the ratings is the work unit, it would be unreasonable to assume that the unique reality of the employee who provides the rating is the only cause of that rating (corresponding to the *fundamental attribution error*; e.g., Ross, 1977). This is not to say that these differences are meaningless, but rather that *climate* constructs reflect employees' exposure to a shared workgroup reality, and that inter-individual deviations should be interpreted as such. These deviations thus reflect either inter-individual differences in employees' level of exposure to specific work characteristics or in their perceptions of these characteristics, and thus cannot be assumed to describe the entirety of individual employees' perceptions of these work characteristics (Marsh et al., 2012; Morin et al., 2014, 2022). This perspective highlights the risk of failing to separate these two layers of influence, especially when focusing on ratings

[1] Contextual constructs are sometimes referred to as reflecting a compilation or formative process of aggregation, whereas climate constructs are sometimes referred to as reflecting a consensus or reflective process or aggregation (Bliese et al., 2007; Klein & Kozlowski, 2000; Lüdtke et al., 2008, 2011; Marsh et al., 2009; Quigley et al., 2007)

explicitly conceptualized, and measured, as characteristics of the work unit. However, this perspective holds irrespective of whether employees are explicitly asked to rate their workgroup reality, or whether they are asked to provide individualized perceptions of their own work reality, as long as these ratings can be expected to capture, to some extent, their shared workgroup reality. However, with the latter option, the multilevel disaggregation will be complicated by the reliance on suboptimal measurement, implicitly forcing employees to incorporate social comparisons to their assessments (i.e., to focus on how they differ from others or on how their reality is unique)[2].

The JD-R model (e.g., Bakker & Demerouti, 2017; Demerouti et al., 2001) is consistent with this perspective, suggesting that exposure to objective or subjective work characteristics should lead to similar effects among employees. For this reason, the JD-R model leads us to expect that job demands and resources should influence employees' levels of need satisfaction in a similar manner at L1 and L2. More precisely, whereas shared perceptions of the job demands and resources present at the work unit level can be expected to influence the extent to which the needs of all team members are met in their workplace, inter-individual differences in these work environment perceptions can also be expected to be associated with increases or decreases in employees' individual levels of need satisfaction beyond that of their teammates. Our methodological approach allows us to test this assumption by verifying whether and how the effects of shared perceptions differ from those of inter-individual variations in these perceptions, while still being able to assume that the referent of those perceptions remains the same (i.e., L2 levels of job demands and resources, or L1 individual variations in the perceptions of the same job demands and resources).

In contrast, variables such as gender, personality, ethnicity, or need satisfaction in the present study, all refer to meaningful individual characteristics of the employees which, when aggregated at the workgroup level, become contextual variables. With contextual constructs, the L1 component is meaningful in its own right in a manner that is independent from the workgroup reality, and the L2 contextual component takes a distinct signification. Here, the question becomes whether this aggregated workgroup reality (e.g., the gender composition of the workgroup or the workgroup-level need satisfaction context) has implications going beyond those of the individual ratings used to create them (e.g., gender or need satisfaction). A good example of contextual effect is the well-established big-fish-little-pond effect from the field of educational psychology (e.g., Marsh et al., 2014a, 2014b). This effect describes how students' academic self-conceptions are impacted by their own academic performance relative to that of their classmates. More precisely, whereas higher achievers tend to display stronger academic self-concepts, exposure to a class of higher achieving students leads to a reduction in students' academic self-concepts irrespective of the effects of individual levels of achievement. Importantly, the unique L2 context formed by the aggregation of workgroup characteristics is also likely to have effects that differ as a function of the outcome considered. For instance, the cultural diversity of a workgroup has been found to benefit creativity (Li et al., 2017), to increase the likelihood of conflict (Jehn et al., 1999), and to have no impact on performance (Webber & Donahue, 2001).

A Need Satisfaction Context

Need satisfaction is conceptualized by SDT as a meaningful individual construct (Ryan & Deci, 2017). Thus, when aggregated at the workgroup level and used as a predictor, need satisfaction becomes a contextual construct. This change in the meaning of aggregated individual variables exists irrespective of whether there is a theoretical rationale to explain whether and how the individual (L1) meaning of need satisfaction differs from that of the workgroup-level (L2) context created by the aggregation of L1 ratings. For instance, there is no need to explain why the gender composition of a workgroup has an effect that deviates from that of gender for this distinction to occur. Yet, acknowledging this phenomenon raises the question of the theoretical meaning taken at L2 by a need satisfaction context formed by the aggregation of individual need satisfaction ratings. Whereas need satisfaction has a clear

[2] We acknowledge the multilevel tradition that contrasts referent-shift (individuals directly rate their group reality) and direct consensus (individuals rate their own reality) approaches to climate ratings (e.g., Wallace et al., 2016), but rather endorse the perspective outlined by Marsh et al. (2012) and Morin et al. (2014, 2022) which highlights the greater theoretical relevance of the former approach to *climate* ratings. Moreover, although consensus approaches have been advocated for ratings of L2 constructs such as the organizational culture, we feel that an assessment of a culture based on a set of shared individual values is closer to a contextual variable or would require procedures more complex than those covered here (Weller, 2007).

meaning at the individual level, a multilevel perspective entails verifying whether its L2 aggregation creates a specific work context playing a truly distinctive role (Kelly et al., 2008).

For instance, if satisfied employees display lower levels of distress, then workgroups including a greater number of satisfied employees should automatically display lower aggregated levels of distress. This workgroup-level effect, however, simply reflects the addition of individual effects. The key question, when contextual effects are considered, is whether combined levels of need satisfaction create a unique context that generates higher, or lower, workgroup levels of psychological distress than would be expected from the simple addition of the individual effects of employees' need satisfaction levels. Competing theoretical perspectives suggest that collective levels of need satisfaction might, indeed, play such a role. However, these perspectives disagree about whether these collective levels of need satisfaction would form a desirable need nurturing context or a more problematic social comparison context.

Need satisfaction is conceptualized by SDT as a fundamental requirement for optimal functioning (Ryan & Deci, 2017). From this perspective, employees whose needs are not met at work can be expected to invest significant efforts at finding ways for their needs to be better fulfilled, sometimes to the expense of other members of their work unit (Van den Broeck et al., 2014). In contrast, employees whose needs are met at work will have more time to support, help, and interact in a positive manner with their colleagues, which may even increase their own level of need satisfaction (Gillet et al., 2020). They will also display more positive attitudes, affects, and behaviors, in addition to feeling less threatened by others given that their own levels of need satisfaction are not threatened at work (Huyghebaert et al., 2022). For these reasons, collective levels of need satisfaction are likely to create a more positive work context for all members of a work unit than could be expected by the simple addition of individual benefits. For instance, members of units characterized by higher aggregated levels of need satisfaction should be more likely to provide each other with constructive but non-threatening feedback aiming at increasing the units' collective mastery rather than on outperforming others, and more frequently encouraged and supported to engage in tasks that are inherently rewarding and aligned with their interests. Likewise, these units might be more frequently characterized by close and authentic relationships.

Supporting these expectations, research has shown that employees characterized by higher levels of need satisfaction are likely to be involved in decision making (Hetland et al., 2011), to take personal initiative and express their true selves (Sonnentag, 2003), to seek out tasks that allow them to develop new skills (Vansteenkiste et al., 2007), to feel free to express their work-related and personal issues (Van den Broeck et al., 2008), and to help each other while showing empathic concern for fellow unit workers (Sheldon, 2005). In contrast, workers characterized by lower global levels of need satisfaction are more prone to jealousy, envy, and competition (Duriez et al., 2007), to experience less authentic relationships (Vansteenkiste et al., 2007), and to feel pressured by ego-involved demands and stressful interpersonal comparisons (Lyubomirsky & Ross, 1997). From this perspective, collective levels of need satisfaction can be expected to create a "need nurturing" context.

In contrast, social comparison processes (Diener & Fujita, 1997; Fujita, 2008; Gerber et al., 2018), may also lead employees to experience negative outcomes from their exposure to a work context characterized by higher collective levels of need satisfaction. Social comparison theory (Gerber et al., 2018) describes the process whereby people rely on others to assess their own situation. This theory highlights the fact that forced (when one does not select with whom to compare oneself, as in the context of an externally defined work unit) upward (when the targets are better off than oneself) comparisons should result in undesirable consequences (Diener & Fujita, 1997; Fujita, 2008; Gerber et al., 2018). Similar to the big-fish-little-pond effect (Marsh et al., 2014a, 2014b), this perspective suggests that, irrespective of their own levels of need satisfaction, employees exposed to highly satisfied work colleagues may come to question their own satisfaction, feeling that they might never be able to attain a level of satisfaction similar to that of their peers, leading them to question their true level of fit within their work unit or organizations. Indeed, upward social comparisons are known to lead employees to feel more uncertain about their abilities, performance, and other socially defined attributes (Lyubomirsky & Ross, 1997). Supporting this assertion, White et al. (2006) demonstrated that people who made more frequent social comparisons were more likely to be unsatisfied with their job, to experience guilt, envy, defensiveness, and regret, to have unmet cravings, to lie, and to blame others. From this perspective, collective levels of need satisfaction can be expected to create a forced upward

social comparison context likely to lead employees to question their place in the workgroup.

The Present Study

The present study provides a first comprehensive assessment of Nielsen et al.'s (2017) multilevel representation of job resources, as well as of the joint JD-R/SDT framework. Although we were limited to the job demands and resources measured as part of the UMP2, the comprehensive nature of this assessment made it possible for us to consider a range of job demands and resources whose role in the prediction of our key outcomes (psychological distress, turnover intentions, and work-to-family conflict) has been established in previous research, at least at the level of individual perceptions[3].

Our first objective is to document the multilevel measurement structure of employees' ratings of the job demands and resources covered in the UMP2. Indeed, despite the conceptually distinct nature of the job demands and resources considered, evidence from previous multilevel research calls into question the true empirical distinctiveness of these dimensions (Demerouti et al., 2001; Gagné et al., 2020; Kiersch & Byrne, 2015). As a result, it would seem appropriate to consider the applicability of alternative bifactor representations of these work characteristics to clearly separate their common core from their unique components. A bifactor solution including four global factors (reflecting job demands, and workgroup, leader, and organization resources) would appear theoretically consistent with Nielsen et al.'s (2017) conceptualization. Alternatively, a solution including two global factors (demands and resources) would be more consistent with the basic formulation of the JD-R model (e.g., Demerouti et al., 2001). Yet, the correlations reported among previous studies also suggest that a third solution including a single global factor, representing employees' holistic perceptions of their work environment quality, might be more suitable.

Research Question 1. *Will employees' ratings of job demands and resources reflect independent dimensions without a common core, or will be better match a bifactor representation including one, two, or four global factors?*

Albeit not a core objective of our study, SDT research has recently converged on the superiority of a bifactor operationalization allowing for the disaggregation of need satisfaction ratings into global and specific components across life domains (e.g., Gillet et al., 2019b), including work (Gillet et al., 2019a, 2020; Sánchez-Oliva et al., 2017). This representation makes it possible to obtain a direct estimate of global levels of need satisfaction, together with a non-redundant estimate of the specific levels of satisfaction of the needs for autonomy, competence, and relatedness left unaccounted for by this global level. This second component directly estimates imbalance in the satisfaction of each need above that of the others. We expect a similar operationalization to be supported across levels.

Hypothesis 1. *Employees' ratings of need satisfaction will follow a bifactor representation including one global factor (global need satisfaction) and three specific factors (autonomy, competence, and relatedness).*

Our second objective is to document the theoretical (predictive) mechanisms underpinning the joint JD-R/SDT framework. These theoretical mechanisms are represented in Figure 1. In this regard, we first consider the likely effect of the job demands and resources covered in the UMP2 in relation to employees' levels of need satisfaction. The bulk of research evidence (see Section 4 of the online supplements), coupled with the generic nature of JD-R predictions (Bakker & Demerouti, 2007, 2017; Demerouti et al., 2001; Nielsen et al., 2017), lead us to expect all job resources to have desirable, and comparable, effects at the individual (inter-individual differences in perceptions) and work unit (aggregated perceptions based on ratings focused on the reality of work unit, thus forming climate constructs) levels, whereas job demands should have comparably undesirable effects at both levels.

Hypothesis 2. *Inter-individual differences in job demands perceptions will be negatively related to employees' levels of need satisfaction, whereas inter-individual differences in job resources perceptions will be positively related to their levels of need satisfaction.*

Hypothesis 3: *Work unit levels of job demands (a climate construct) will be negatively related to work unit levels of need satisfaction, whereas work unit levels of job resources (a climate construct) will be positively related to work unit levels of need satisfaction.*

---

[3] We provide a substantive real-data application of the core principles of doubly-latent multilevel estimation outlined by Morin et al. (2022). In Sections 1 to 3 of the online supplements, we summarize the methodological issues outlined by Morin et al. (2022). Then in Section 4, we review previous multilevel JD-R research, with a particular attention to whether previous studies have attended to the considerations outlined in Sections 1-3.

Our third objective is to test the role of psychological need satisfaction as a mediator of the effects of job demands and resources on a series of outcomes reflecting employees' turnover intentions, psychological distress, and work-to-family conflict, as expected in the joint JD-R/SDT framework (Boudrias et al., 2011, 2011; Van den Broeck et al., 2008). Whereas this proposition has been supported in single-level analyses, no previous study has considered these associations via a proper multilevel disaggregation of the effects occurring at the individual and work unit levels. Given our divergent theoretical expectations regarding the possible benefits (need nurturing context) or risks (forced upward social comparisons) associated with the context formed by need satisfaction at the work unit level, and the fact that these effects are likely to change or to generalize across outcomes, we consider the nature of these effects as an open research question.

Hypothesis 4. *Individual levels of need satisfaction will be negatively related to turnover intentions, psychological distress, and work-to-family conflict.*

Hypothesis 5. *Individual levels of need satisfaction will mediate the effects of job demands and resources on turnover intentions, psychological distress, and work-to-family conflict.*

Hypothesis 6: *Work unit levels of need satisfaction (a contextual construct) will mediate the effects of work unit levels of job demands and resources on work unit levels of turnover intentions, psychological distress, and work-to-family conflict.*

Research Question 2. *Will the need satisfaction work context be associated with additional benefits or risks, and will these effects vary across outcomes?*

Method

Participants and Procedures

In this study, we rely on data collected by CAF/DND among a sample of 5,716 military and civilian personnel who completed the UMP2 assessment (Ivey et al., 2018). The UMP2 assessment was approved by the CAF/DND Social Science Research Review Board. All respondents provided informed consent and were guaranteed that their individual responses would be kept confidential, with reporting based solely on aggregated data. Participants were recruited in 50 distinct work units, comprising an average of 114.32 employees (range: 15 to 358; *SD* = 80.55). The average response rate across work units was 68%. Most participants (76%) completed the English version of the questionnaires, whereas the others completed the French version. Most respondents were military (72%), 15% were civilian, and 13% did not report their status, whereas 66% of the respondents were males, 18% were females, and the remainder (16%) did not indicate their gender.

Measures

Measures (supervisor safety actions and expectations, role ambiguity, work overload, and turnover intentions) not already validated in French, were adapted by professional translators from the Government of Canada's Translation Bureau. These items were then back-translated into English by bilingual CAF/DND experts. Discrepancies were resolved by consensus. When completing all measures related to job demands and resources, employees were asked to rate the reality of their work unit rather than their own individual reality.

Team psychological safety (workgroup resource). Team psychological safety was assessed using a seven-item scale (α = .84; e.g., *It is safe to take a risk on this team*) developed by Edmondson (1999; French by Martins et al., 2012). Participants rated whether the items were an accurate descriptor of their team reality using a seven-point scale (1-*Very Inaccurate* to 7-*Very Accurate*).

Interpersonal justice (workgroup resource). Interpersonal justice perceptions were measured with the four-item subscale (α = .93; e.g., *Refrain from improper remarks or comments*) from Colquitt's (2001; French by Camerman et al., 2007) questionnaire. Items followed the stem "*Please indicate the extent to which individuals (coworkers, supervisors, etc.)…"* and were rated on a five-point scale (1-*To a Very Small Extent* to 5-*To a Very Large Extent*) in relation to the work unit.

Contingent recognition and reward (supervisor resources). Perceptions of supervisors' contingent recognition and reward were measured with the relevant four-item subscale (α = .85; e.g., *There are few rewards for those who work here* – all items reversed coded) of the Job Satisfaction Survey (Spector, 1985; French by Lapierre, 2013). All items were rated in reference to supervisors' practices using a six-point scale (1-*Disagree Very Much* to 6-*Agree Very Much*).

Transformational leadership (supervisor resource). Perceptions of transformational leadership were assessed using the seven-item (α = .96; e.g., *Communicates a clear and positive vision of the future*) Global Transformational Leadership Scale (Carless et al., 2000; French by Fernet et al., 2015).

Items followed the stem "*Please indicate how often your supervisor…*" on a five-point frequency scale (1-*Rarely or Never* to 5-*Very Frequently, if not Always*).

Supervisor's safety actions and expectations (supervisor resources). Perceptions of the extent to which supervisors' actions (five items; α = .71; e.g., *My supervisor says a good word whenever he/she sees a job done according to the safety rules*) and expectations (five items; α = .84; e.g., *Whenever pressure builds up, my supervisor wants us to work faster, rather than by the rules*– all items reversed coded) sought to protect physical safety were measured with Zohar's (2000) scales. Items were rated on a five-point scale (1-*Completely Disagree* to 5-*Completely Agree*).

Organizational support (organizational resource). Perceptions of organizational support were assessed using the four-item (α = .88; e.g., *The organization really cares about our well-being*) version (Caesens et al., 2014, who also provided the French version) of Eisenberger et al.'s (1986) measure. These items were rated on a seven-point response scale (1-*Strongly Disagree* to 7-*Strongly Agree*).

Culture of interpersonal respect (organizational resource). Perceptions of the extent to which the organizational culture values interpersonal respect were measured with three items (α = .87; e.g., *Respect for the individual's rights*) from the Organizational Culture Profile (O'Reilly et al., 1991; French by Vandenberghe, 1999). Items followed the stem "*Please indicate to what extent each of the following characteristics describe the culture of your organization*" and were rated on a five-point scale (1- *Not at All* to 5-*To a Great Extent*).

Role ambiguity (job demand). Role ambiguity was assessed with six items (α = .91; e.g., *We know what our responsibilities are* – all items reversed coded; Gonzàlez-Romà & Lloret, 1998) rated in relation to the work unit reality on a seven-point scale (1-*Strongly Disagree* to 7-*Strongly Agree*).

Work overload (job demand). Work overload was assessed with the six-item (α = .91) short version (Thiagarajan et al., 2006) of Reilly's (1982) questionnaire. Items (e.g., *We need more hours in the day to do all the things that are expected of us*) were rated on a seven-point scale (1-*Never* to 7-*Always*).

Need satisfaction (mediator). Participants indicated if they felt that their individual needs for competence (four items; α = .89; e.g., *I am good at the things I do in my job*), relatedness (six items; α = .88; e.g., *At work, I feel part of a group*), and autonomy (six items; α = .81; e.g., *I feel free to do my job the way I think it could best be done*) were met at work with the Work-Related Basic Needs Satisfaction Scale (Van den Broeck et al., 2010; French by Gillet et al., 2020). Items were rated on a five-point scale (1-*Totally Disagree* to 5-*Totally Agree*).

Outcomes. Kessler et al.'s (2002) Psychological Distress Scale (French by Arnaud et al., 2010) was used (10 items; α = .93; e.g., *Did you feel depressed?*). All items were presented after the stem "*In the last four weeks, about how often…"* and rated on a five-point scale (1-*None of the Time* to 5-*All the Time)*. Work-to-family conflict was measured with five items (α = .96; e.g., *The demands of my work interfere with my home and family life;* Netemeyer et al., 1996; French by Simon et al., 2004), rated on a seven-point response scale (1-*Strongly Disagree* to 7-*Strongly Agree*). Turnover intentions were assessed with three items (α = .82; e.g., *I frequently think of quitting my job*; Colarelli, 1984) rated on a five-point scale (1-*Strongly Disagree* to 5-*Strongly Agree*).

Analyses

Mplus 8.8 (Muthén & Muthén, 2022) was utilized to conduct all analyses via the Maximum Likelihood-Robust (MLR) estimator, which is robust to multilevel nesting and non-normality. Missing data were handled with Full information Maximum Likelihood (Enders, 2010). Our main models were specified as multilevel with employees (L1) nested under work units (L2). Preliminary measurement models were estimated at the individual level relying on Mplus design-based correction procedures (Asparouhov, 2005) to control for the multilevel data structure. Model fit was estimated and contrasted using the Root Mean Square Error of Approximation (RMSEA), the Tucker-Lewis Index (TLI), and the Comparative Fit Index (CFI). The robust χ^2 will also be reported. According to common guidelines, RMSEA values under .06 and .08, and TLI/CFI values above than .95 and .90 respectively support excellent and acceptable fit (Hu & Bentler, 1999; Marsh et al., 2004, 2005).

Measurement Models (Research Question 1, Hypothesis 1)

Single-level CFA and DL-MLCFA were used to assess the a priori measurement structure of all constructs, as well as their isomorphism (i.e., metric invariance, equal factor loadings) across levels (necessary to ensure comparability in meaning across levels; Morin et al., 2022). For more details about critical multilevel measurement considerations, see Section 1 of the online supplements. These verifications were conducted on a single type of construct at a time (work characteristics, need

satisfaction, and outcomes). The outcomes model was the simplest, involving the estimation of three correlated CFA factors (reflecting turnover intentions, psychological distress, and work-to-family conflict) at both levels. Two a priori L1 correlated uniquenesses were integrated to account for the parallel wording of two pairs of psychological distress indicators (Marsh et al., 2013). For need satisfaction, we contrasted a three-factor (relatedness, competence, and autonomy) CFA model with the a priori bifactor[4] solution advocated in previous research (Gillet et al., 2019a, 2020; Sánchez-Oliva et al., 2017). One a priori L1 method factor was integrated to account for the negative wording of six need satisfaction items (Marsh et al., 2013).

A more comprehensive examination was needed for work environment characteristics. In a first model, ten correlated factors representing role ambiguity, work overload, team psychological safety, team interpersonal justice, contingent recognition and reward, transformational leadership, supervisor safety actions, supervisor safety expectations, organizational support, and organizational culture of respect were estimated. In a second model, a bifactor approach allowed all work environment items to define a single global work environment factor (G-factor) together with the ten a priori specific factors (S-factors). In a third model, an alternative bifactor approach involving two correlated G-factors reflecting job demands and job resources were estimated jointly with the ten S-factors. Finally, a fourth model relied on bifactor alternative including four correlated G-factors reflecting job demands, team resources, supervisor resources, and organizational resources, together with ten S-factors. All models include one correlated uniqueness to account for the negative wording of one pair of organizational support items (Marsh et al., 2013). Once the optimal model for all constructs was identified, a single global measurement model was estimated.

We relied on these measurement models to estimate factor correlations, inter-rater agreement via the calculation of two intraclass correlation coefficients (ICC1 and ICC2), and composite reliability via the calculation of McDonald's (1970) omega (ω, ω_{L1}, and ω_{L2}). DL-MLCFA and DL-MLSEM are explicitly corrected for all three sources of unreliability (inter-item reliability at L1; inter-item reliability at L2; and inter-rater reliability when estimating the L2 constructs). For more details about reliability, see Section 2 of the online supplements. We also conducted tests of measurement invariance to ascertain that responses were equivalent across samples of English and French, males and females, and civilian and military participants (Millsap, 2011). These tests were conducted in the following sequence: i) configural invariance (same model); ii) weak invariance (same factor loadings); iii) strong invariance (same item intercepts); iv) strict invariance (same item uniquenesses); v) correlated uniquenesses invariance; vi) latent variance-covariance invariance; and vii) latent means invariance. These measurement invariance results are reported in Table S5 of the online supplements, and support the equivalence of all measures across language, sex, and military or civilian status.

Predictive Models (Hypotheses 2 to 6, Research Question 2)

The global DL-MLCFA solution was used as the basis for our a priori DL-MLSEM model. This model assumed matching relations among constructs at both levels. For a more extensive discussion of the implications of climate (work characteristics) or contextual (need satisfaction) constructs for centering, see Section 3 of the online supplements. Because our main focus was on the associations involving the global (G-factors) work environment and need satisfaction constructs, our a priori model (P1) allowed the work environment G-factor(s) to predict the need satisfaction G-factor, which in turn was allowed to predict the outcomes (work-to-family conflict, turnover intentions, and psychological distress). We contrasted this fully mediated model (Preacher et al., 2010) to a partially mediated model (P2) in which the work environment G-factor(s) was(were) allowed to predict the outcomes. Alternative models were then estimated to verify the role of the S-factors: i) the work environment S-factors were allowed to predict the need satisfaction G-factor (P3); ii) the work environment G-factor(s) was allowed to the predict the need satisfaction S-factors (P4); iii) the work environment S-factors were allowed to predict the need satisfaction S-factors (P5); iv) the need satisfaction S-factors were allowed to predict the outcomes (P6); and v) the work environment S-factors were allowed to predict the outcomes (P7).

Mplus reports standardized coefficients separately at each level, which is not appropriate for DL-

[4] Bifactor models (see also Section 1 of the online supplements) disaggregate ratings of multiple dimensions into non-overlapping (uncorrelated) global (G-factor) and specific (S-factors) components. The G-factor reflects the variance shared across all items, whereas the orthogonal S-factors reflect the subscale-specific variance left unexplained by the G-factor (reflecting deviations from the G-factor (Morin et al., 2017, 2020).

MLSEM given that multilevel indirect effects occur across levels (Preacher et al., 2010), and that contextual effects are calculated as differences between the L2 and L1 effects (Marsh et al., 2012, Morin et al., 2014). Following formulas provided by Marsh et al. (2009), we standardized coefficients (β) in relation to the total (L1 and L2) outcome variance ($\beta = b * SD_{predictor})/SD_{outcome}$), and calculated effect sizes (ES) based on the L1 outcome variance ($ES = b*SD_{predictor})/SD_{outcomeL1}$). These coefficients (β and ES), as well as all contextual (reflecting L2 effects of need satisfaction on the outcomes) and indirect effects were calculated using Mplus' MODEL CONSTRAINT. To assess the statistical significance of indirect effects, we relied on DL-MLSEM Monte Carlo confidence intervals (Bauer et al., 2006; Preacher & Selig, 2012) calculated using 20,000 replications using Preacher and Selig's (2012) online calculator (quantpsy.org). The Mplus input used to estimate the final DL-MLSEM, as well as a Figure (S1) representing this model, are provided at the end of the online supplements.

Results

Measurement Models (Preliminary Analyses, Research Question 1, and Hypothesis 1)

Outcomes. The fit of all models is reported in Table 2. These results supported the adequacy of the outcomes measurement model. This model had an acceptable level of fit to the data in the single-level and multilevel models, and demonstrated cross-level isomorphism (i.e., the more parsimonious model in which loadings were set to be equal across levels resulted in an equivalent level of fit relative to the model in which these loadings were free to vary across levels). Parameter estimates from this model are reported in Table S1 of the online supplements, while reliability information is reported in Table 3. These results reveal factors defined by strong factor loadings for psychological distress (λ= .647 to .988 across levels and solutions, M_λ = .801; ω = .930; ω_{L1}= .928; ω_{L2} = .979), turnover intentions (.823 to .997, M_λ = .927; ω = .825; ω_{L1}= .826; ω_{L2} = .729), and work-to-family conflict (.588 to .864, M_λ = .748; ω = .959; ω_{L1}= .949; ω_{L2} = .994).

Need Satisfaction. For the need satisfaction models, the a priori bifactor model had a substantially better fit than its first-order counterpart in single-level and multilevel solutions. The multilevel bifactor model was also isomorphic across levels (i.e., equivalent fit). The parameter estimates from this model are reported in Table S2 of the online supplements[5], while reliability information is reported in Table 3. Parameter estimates revealed a well-defined G-factor (.226 to .924, M_λ = .498; ω = .886; ω_{L1}= .881; ω_{L2} = .957), although the G-factor loadings of the competence items (.264 to .447, M_λ = .328) were slightly lower than those of the relatedness (.248 to .924, M_λ = .584) and autonomy (.226 to .836, M_λ = .525) items. This G-factor was accompanied by well-defined relatedness (.266 to .604, M_λ = .488; ω = .809; ω_{L1}= .804; ω_{L2} = .792), autonomy (.126 to .841, M_λ = .445; ω = .626; ω_{L1}= .619; ω_{L2} = .870) and competence (.716 to .897, M_λ = .783; ω = .878; ω_{L1}= .876; ω_{L2} = .943) S-factors. These results thus support Hypothesis 1.

Work Environment. The single-level work environment models all had an acceptable level of fit, although the fit of the bifactor models was substantially higher than that of the first-order CFA. The fit of the bifactor solution incorporating four G-factors was substantially higher than that of the alternative models. The pattern of results was similar for the multilevel models once isomorphism was imposed, as most solutions failed to achieve a satisfactory level of fit without imposing isomorphism, which resulted in an improvement in model fit for most models and was thus supported by the data.[6]

---

[5] When interpreting a bifactor solution, one should keep in mind that bifactor models divide the reliable (i.e., true score) item variance into two components (the G- and S-factors), leading to factor loadings and reliability estimates that tend to be smaller than those observed in first-order models (Morin et al., 2020). This is why it has been argued that more lenient guidelines should be applied when considering the reliability of S-factors from bifactor models (approaching ≥ .50, rather than .60 or .70). Morin et al. (2020) note that well-defined G- and S-factors should ideally be accompanied by large enough (minimally higher than .300, but ideally higher than .500 on at least one of the G- or S-factors) to support their interpretation as key indicators. Indeed, bifactor solutions frequently reveal items that are predominantly associated with only one of these two sets of factors. Such observations do not indicate problems with these items, but rather demonstrate that they represent stronger indicators of one layer of measurement (Morin et al., 2020). Finally, despite the fact that support for a bifactor solution requires that at least some S-factors retain some specificity, bifactor solutions are known to be robust to vanishing S-factors, which simply suggest that the items associated with these subscales mainly serve to define the G-factor and retain limited specificity once this G-factor has been taken into account (Morin et al., 2020).

[6] This observation is consistent with the results from Lüdtke et al.'s (2008, 2011) simulations studies, who showed that isomorphism had a stabilizing impact on model estimation.

These solutions revealed correlations as high as .827 at L1 and .961 at L2 in the first-order CFA, consistent with one or more G-factors and with previous research evidence suggesting conceptual overlap at L2, although this overlap might have been missed in single-level models. In the bifactor solution including four G-factors (i.e., the best fitting solution), the G-factor correlations were again high enough to suggest conceptual overlap (L1 $|r|$ = .739 to .869; L2 $|r|$ = .659 to 1.011). The bifactor solution including two G-factors (demands and resources) similarly suggested conceptual redundancies (L1 r = -.897; L2 r = -1.031). The bifactor solution including a single work environment G-factor was thus retained in response to our Research Question 1. The results from this model are reported in Table S3 of the online supplements, while reliability information is reported in Table 3.

These results first revealed a well-defined G-factor with positive loadings from most of the job resources indicators across models and levels (.063 to .960, M_λ = .643) and negative loadings from the job demands indicators (-.164 to -.933, M_λ = -.484). This G-factor was also associated with a composite reliability (ω = .976; ω_{L1}= .974; ω_{L2} = .994). Most S-factors retained a meaningful level of specificity and reliability: Interpersonal justice (.451 to .696, M_λ = .583; ω = .874; ω_{L1}= .870; ω_{L2} = .834), transformational leadership (.338 to .610, M_λ = .509; ω = .918; ω_{L1}= .916; ω_{L2} = .973), supervisor safety actions (.375 to .943, M_λ = .560; ω = .641; ω_{L1}= .613; ω_{L2} = .964), supervisor safety expectations (.362 to .675, M_λ = .524; ω = .762; ω_{L1}= .760; ω_{L2} = .789), culture of interpersonal respect (.357 to .575, M_λ = .475; ω = .689; ω_{L1}= .684; ω_{L2} = .906), role ambiguity (.177 to .720, M_λ = .444; ω = .808; ω_{L1}= .808; ω_{L2} = .727), and work overload (.579 to .981, M_λ = .777; ω = .904; ω_{L1}= .901; ω_{L2} = .970). The team psychological safety (.244 to .507, M_λ = .340; ω = .550; ω_{L1}= .543; ω_{L2} = .850), contingent recognition and reward (.128 to .642, M_λ = .400; ω = .677; ω_{L1}= .674; ω_{L2} = .803), and organizational support (.235 to .569, M_λ = .401; ω = .586; ω_{L1}= .578; ω_{L2} = .925) S-factors did not retain as much specificity, suggesting that items from these dimensions played a stronger role in defining the G-factor than their S-factors. Interestingly, the S-factors with the lowest reliability at L1 (i.e., supervisor safety actions, culture of interpersonal respect, team psychological safety, contingent recognition and reward, and organizational support) all displayed a much higher level of reliability at L2, consistent with their nature as climate constructs. However, results associated with these S-factors at L1 should thus be interpreted with caution.

Global Model and Intraclass Correlations. These three models were combined in a global measurement model, which had an acceptable level of fit to the data (Table 2). This model was retained for our main analyses. The L1 and L2 factor correlations from this model are reported in Table S4 of the online supplements and are consistent with relations involving mainly the G-factors.

Intraclass correlations calculated from this final model are reported in Table 3. In terms of inter-rater reliability (agreement in the ratings of the L2 constructs), the ICC1 revealed that both G-factors (.066 for work environment and .058 for need satisfaction), all outcomes (.035 for psychological distress, .045 for turnover intentions, and .188 for work-to-family conflict), and most S-factors (.039 to .123 with the exception of .016 for psychological safety, .018 for transformational leadership, .015 for safety actions, .022 for role ambiguity, and .018 for relatedness) had enough between-group variance to support multilevel analyses, especially with the large sample available here. Furthermore, the inter-rater reliability of all L2 constructs was satisfactory (ICC2 = .634 to .964, M = .817), particularly when relying on models providing a correction for this source of measurement errors.

Predictive Analyses (Hypotheses 2 to 6 and Research Question 2)

The fit from the predictive models is reported in the bottom of Table 2. These results support the a priori predictive model (P1) and showed that associations were limited to the global factors. Not only was this model able to achieve acceptable fit, it also achieved a level of fit that was either equivalent or better than the less parsimonious alternatives (P2 to P7). Parameter estimates from the final statistical model estimated in this study are reported in Table 4. At the individual level (L1), the results revealed a strong positive association between inter-individual differences in global work environment perceptions and employees' global need satisfaction, coupled with moderate to strong negative relations between global need satisfaction and turnover intentions, psychological distress, and work-to-family conflict. These results also confirmed negative indirect associations between inter-individual differences in global work environment perceptions and the outcomes, as mediated by global need satisfaction. Thus, although our results did not allow us to separate the effects of job demands from those of job resources, our results are generally consistent with Hypotheses 2, 4, and 5.

Raw L2 results matched their L1 counterpart but were slightly weaker when standardized

coefficients and ES indicators were considered. These results showed a moderate positive association between the global quality of the work units' environment (a climate effect) and aggregated levels of global need satisfaction, thus generally supporting Hypothesis 3. L1 relations between global levels of need satisfaction and the outcomes generalized to the work unit level due to the aggregation of individual-level effects, although the magnitude of these relations was diluted. These three associations resulted in statistically significant L2 indirect associations between work units' global work environment quality and the outcomes, as mediated by aggregated global levels of need satisfaction, providing early support for Hypothesis 6 pending the verification of contextual effects.

When L2 contextual effects were considered, responses to our Research Question 2 differed as a function of the outcome considered. First, work units' global levels of need satisfaction did not create a context leading to additional effects on psychological distress beyond the simple combination of individual-level effects. However, work units' global levels of need satisfaction did carry contextual effects related to the prediction of turnover intentions and work-to-family conflict. Exposure to a work context characterized by higher global levels of need satisfaction contributed to a reduction in work-to-family conflict beyond the negative effects of employee's global levels of need satisfaction (a positive need nurturing effect). However, these contextual effects were not without cost, as they also contributed to increased turnover intentions at the work unit level (a negative social comparison effect), thus limiting the benefits of individual employees' global levels of need satisfaction. This contextual effect resulted in statistically significant L2 indirect associations between work units' global work environment quality and these two outcomes (a positive indirect effect for turnover intentions and a negative one for work-to-family conflict), as mediated by the global need satisfaction context, thus partially supporting Hypothesis 6 in relation to two, out of three, outcomes.

Discussion

The results from this study have broad implications for research focusing on the effects of work environment characteristics on employees, and particularly for research anchored in the JD-R model and SDT. Many of these implications would have been impossible to establish without DL-MLSEM.

A Global Perspective on the Measurement and Effects of Job Demands and Resources

Answering the call for multilevel JD-R research (e.g., Bakker & Demerouti, 2017; Nielsen et al., 2017), this study is the first to demonstrate that employees' ratings of a variety of job demands and resources followed a similar bifactor representation at the individual and work unit levels. More precisely, across levels, employees' perceptions of the job demands and resources present in their work environment were first captured by one global factor reflecting the overarching quality of that work environment in terms of providing resources allowing employees to meet their work demands. This global work environment factor proved to be a reliable indicator of employees' individual global perceptions of their work environment quality (ω_{L1} = .974) as well as of work units' aggregated work environment quality (ω_{L2} = .994; ICC2 = .889). However, most of the job demands and resources covered in this study retained something uniquely distinct from this global factor. Even though some of these characteristics (i.e., team psychological safety, contingent recognition and reward, and organizational support) retained less specificity than others (i.e., interpersonal justice, transformational leadership, supervisor safety actions and expectations, culture of interpersonal respect, role ambiguity, and work overload) it was still impressive to note that the all of them proved to be reliable indicators of the work unit reality to which employees felt being exposed (ω_{L2} = .727 to .973; ICC2 = .634 to .942). Thus, it is not that employees necessarily perceived their work environment characteristics as either being all "good" or all "bad", but rather that they are able to hold reasonably distinct perceptions of these characteristics, and that these perceptions still converge to some extent toward the assessment of one indicator of their global quality of the work environment.

However, our results also highlighted that predictions were limited to this global indicator of work environment quality, revealing no residual effects associated with specific perceptions of job demands and resources beyond these global perceptions. This gestalt of employees' work environment perceptions thus seems to represent a far more important driver of need satisfaction than their perceptions of any specific job demand or resource. This does not mean that specific job demands and resources are not important. Indeed, ratings of all of these characteristics were used in the estimation of this global work environment dimension, so that taking any of these out is likely to change the nature of the global construct. For the moment, and pending replication, this G-factor should be seen as a synthesis of all variables considered and suggests that the effects of strengths or limitations associated

with specific work characteristics (e.g., transformational leadership, work overload) can be offset by other characteristics, as long as global perceptions remain sufficiently positive. What remains unknown, however, is whether these compensatory mechanisms are conditioned by any kind of turning point (i.e., is there a threshold under which positive perceptions of work environment characteristics are unable to compensate for negative perceptions of other characteristics?), or whether this global indicator truly reflects a gestalt of all characteristics captured by the model. Likewise, it remains unknown whether all job characteristics play a similar compensatory role, or whether some might be more potent drivers of global perceptions. The present study also solely focused on work-related resources, neglecting the potentially critical role of individual resources (e.g., Xanthopoulou et al., 2007). This limitation might be important given previous reports showing that some individual resources, such as job crafting, might help employees create a more positive work environment for themselves (Sánchez-Cardona et al., 2020). Moreover, different employees may also perceive similar job demands differently as either a challenge or a hindrance based on their individual characteristics, and these perceptions are also likely to modify their reactions to these characteristics (Li et al., 2020).

In sum, we initially described the JD-R model as a broad heuristic likely to help researchers anticipate the effects of any type of job characteristics based on its classification as a job demand or resource. In alignment with this broadband perspective, our results demonstrated that ratings obtained across a wide range of job demands and resources did indeed capture a common core. They also showed that, even though employees can reliably differentiate distinct types of job demands and resources, this common core remained the core driver of associations. Our results also highlighted that the bifactor approach to measurement advocated in this study might be able to capture this common core in a way that is somehow independent from the measures used to assess it. However, this assumption needs to be thoroughly investigated in the context of research focusing on distinctive sets of job demands and resources and considering their role for an even wider range of outcomes while also accounting for inter-individual differences in perceptions. However, and pending further investigation, our results suggest that managerial assessments and interventions should focus on a broad range of characteristics to achieve such an overarching picture of the workplace.

A Multilevel JD-R/SDT Perspective

Our research also sought to validate the joint JD-R/SDT framework according to which the effects of work characteristics should be mediated by employees' level of need satisfaction (Van den Broeck et al., 2008, 2016). To verify the possible contextual nature of employee's need satisfaction ratings, we focused on relations well-documented at the individual level (Boudrias et al., 2011, 2014; Deci et al., 2017; Gillet et al., 2015; Van den Broeck et al., 2016) but insufficiently documented at the work unit level. The results supported our a priori predictive model at the individual level, thus validating the role of need satisfaction as a mediator of the effects of inter-individual differences in global work environment perceptions on employees' turnover intentions, psychological distress, and work-to-family conflict. All of these associations were also apparent at the work unit level (simply reflecting the sum of individual effects), albeit slightly diluted.

More precisely, our findings first replicated previous results (Gillet et al., 2019a, 2020; Sánchez-Oliva et al., 2017) supporting the superiority of a bifactor operationalization of need satisfaction, while demonstrating for the first time the generalizability of this structure across the individual and work unit levels. In this representation, a reliable global need satisfaction construct (ω_{L1} = .881; ω_{L2} = .957; ICC2 = .875) was found to reflect the extent to which the satisfaction of all three needs is aligned, whereas specific factors were found to reflect the imbalance in the satisfaction of each need beyond this global level (which retained a substantial amount of specificity beyond the global factor). This operationalization is consistent with Sheldon and Niemiec's (2006) proposition that a balanced level of need satisfaction (captured in this study by our global factor) was essential for well-being, as well as with Gillet et al.'s (2019a, 2020) demonstration that a bifactor approach was a way to directly assess individuals' global levels of need satisfaction across all three needs, together with direct and non-redundant estimate of the degree of imbalance in the satisfaction of each specific need. This operationalization suggests that the only way to increase this global level of need satisfaction is through improving the satisfaction of all three needs in a balanced manner, as the imbalanced satisfaction of any single need is rather likely to be absorbed by the specific factors.

Second, our results showed that, across levels, all associations were limited to employees' global levels of need satisfaction (with no residual effects associated with imbalance in the satisfaction of each

specific need), which was found to completely mediate the effects of work environment characteristics across all outcomes considered in this study. Thus, although imbalanced levels of need satisfaction were not found to play a role in the present study, the observation that the effects of need satisfaction were limited to this global level remains consistent with the theoretical perspective that a balanced level of need satisfaction across all three needs represents a core driver of associations between work environment characteristics and a variety of outcomes (Sheldon & Niemiec, 2006). Clearly, additional research is needed to document the situations under which imbalance in specific need satisfaction might also play a role. However, although the interpretation of the effects of global levels of need satisfaction occurring at the individual level are fairly straightforward, that of the need satisfaction work context occurring at the work unit level require some further elaboration.

An Outcome-Specific Need Satisfaction Context

Turning our attention to the properly calculated contextual effects of need satisfaction, our results suggest that, at the work unit level, need satisfaction creates a context whose effects differ across outcomes. On the one hand, we found no evidence that need satisfaction yielded any additional contextual effects in the prediction of psychological distress relative to aggregated individual effects. This lack of contextual effect does not imply that global levels of need satisfaction have no L2 effect. Rather, it simply shows that these global levels do not lead to the formation of a specific workgroup context that, in and of itself, impacts workgroup levels of psychological distress beyond the effects already created by the aggregated levels of global need satisfaction of workgroup members. Such contextual effects would indicate that exposure to co-workers with high levels of global need satisfaction resulted in additional effects beyond the effects of employees' individual levels of global need satisfaction. This did not happen in this study as far as psychological distress was considered, which could be related to the domain-general nature (not work-specific) of our measure of psychological distress.

On the other hand, work unit's global levels of need satisfaction carried contextual effects related to turnover intentions and work-to-family conflict. In relation to work-to-family conflict, exposure to a work context characterized by higher global levels of need satisfaction contributed to a reduction of work-to-family conflict beyond the negative effects of aggregated global levels of need satisfaction. This result supported our expectations in showing that exposure to highly satisfied colleagues seemed to create a positive need nurturing work context. This context may help to demonstrate to employees that their needs are important for the organization, which in turn may make it easier for them to limit the extent to which work demands are allowed to interfere with their family life. As noted in the introduction, satisfied workers share more positive social relationships in their work unit (Gillet et al., 2020), more frequently display interpersonal citizenship behaviors such as altruism and helping (Huyghebaert et al., 2022), take more initiative, and express their opinions more freely (Hetland et al., 2011; Van den Broeck et al., 2008). At the work unit level, these characteristics are likely to benefit employees exposed to more satisfied colleagues, thus creating a work environment characterized by closer and more authentic workplace relationships, as well as freedom and independence, which are all likely to limit the emergence of work-to-family conflict (Weale et al., 2019).

In contrast, the same work context resulted in a slight increase in turnover intentions, over and above the effects of individual levels of need satisfaction on decreases in turnover intentions. As suggested in our introduction, this effect could potentially be explained by exposing employees to forced (i.e., one does not select one's workmates) upward (i.e., involving highly satisfied colleagues) social comparison processes (e.g., Diener & Fujita, 1997; Fujita, 2008; Gerber et al., 2018). This exposure may lead employees to question their own satisfaction, leading them to interrogate their true desire to stay in an organization where they are unlikely to ever experience a level of satisfaction matching that of their peers. Our results suggest that similar processes seemed to be involved when explaining the relations between the need satisfaction context and turnover intentions. An alternative, and possibly complementary, explanation for this finding is that the need satisfaction work context could also lead employees to feel more supported in living their own dreams and pursuing their own aspirations, possibly within other organizations. This interpretation is consistent with prior research showing that need satisfaction is associated with higher levels of career satisfaction and commitment as well as perceived person-vocation fit (Dahling & Lauricella, 2017; Onyishi et al., 2019).

Limitations and Directions for Future Research

Despite the strengths of this study, it still presents limitations. First, it remains cross-sectional,

making it impossible to reach conclusions on the directionality of the associations. Although directionality was theoretically grounded, this grounding is not sufficient to overshadow the benefits of longitudinal research. Longitudinal research provides a way to assess directionality, as well as stability, change, and trajectories. In addition, calls have been made to increase researchers' focus on time-sensitive within-person fluctuations in need satisfaction levels (e.g., Hooff & De Pater, 2019), which would be worth investigating with multilevel considerations, allowing one to study how work unit characteristics impact individual trajectories. Second, despite relying on a large multilevel sample and finding support for the generalizability of the measurement models across males and females, military and civilian, and French and English participants, the generalizability of our results outside of the Canadian Defence context remains unknown. Still, it is important to note that our results support the value of the UMP2 for the assessment of work units, as workgroup-level effects remained significant. From a theoretical perspective, it is however important to reinforce that the UMP2 was primarily, albeit not exclusively, driven by practical, rather than theoretical, considerations. As a result, it would be easy to argue, from a theoretical standpoint, that some critical job demands and resources should have been included to this assessment process or that some of those that were included might not have been the most critical ones to consider. Importantly, although the UMP2 differentially assessed job resources related to workgroup, leader, and organization anchored in Nielsen et al.' (2017) multilevel conceptualization, job demands are equally likely to share a similar (workgroup, leader, organization) structure that has yet to be systematically conceptualized or validated. As a result, our operationalization of the work environment has a much stronger focus on job resources than on job demands, which might have played a role in our results. These observations reinforce the need for theoretically-driven replication studies, but also highlights the interest of the bifactor approach adopted in the present study. Indeed, if this approach was found to generalize to other test batteries, organizational contexts, cultures, and operationalizations of the job demands and resources, it would suggest that conclusions, to some extent, generalize to different sets of job demands and resources.

Third, although our results showed a lack of contextual effects of need satisfaction on psychological distress, these results are limited to one of three outcomes. Psychological distress presented a limited level of between-group variability (ICC1 = .035) which might have made it harder to detect contextual effects. The presence of raw L2 effects reinforces the idea that this smaller between-group variance was not enough to preclude the detection of L2 effects. Yet, it also reinforces the need to replicate our results with a broad range of outcomes presenting more between-group variability, or directly measured at L2 (e.g., official ratings of unit-level absenteeism). Another limitation stems from our reliance on self-reports. Fortunately, as shown in Siemsen et al.'s (2010) mathematical demonstration, shared method variance cannot play a role in multivariate analyses, in addition to being naturally controlled for in DL-MLSEM (where group-mean centering re-expresses scores as deviations from the group mean). Yet, it would have been informative to combine these measures with more objective measures of workplace functioning (e.g., observational measures or supervisors' reports), psychological health (e.g., medical leaves), or work behaviors (e.g., absenteeism, turnover). More generally, despite our assertion (e.g., Marsh et al., 2012; Morin et al., 2014, 2022) that climate constructs are best captured by referent-shift approach (asking employees to rate their workgroup reality), it would be interesting to contrast the results obtained via this approach relative to that relying on a consensus approach (individuals rate their own reality; e.g., Wallace et al., 2016).

Lastly, for researchers considering implementing these models to their research, one final consideration remains important: Sample size. Although this consideration was not directly relevant in the present study due to the availability of a reasonably large sample, it is likely to be far more relevant for applications relying on more modest samples. Interested readers are referred to Section 6 of the online supplements for a more extensive presentation of these considerations.

Practical Implications

From an applied perspective, changes designed to increase job resources and/or to decrease job demands at both the individual and work unit levels might prove fruitful from the perspective of supporting the satisfaction of employees' basic psychological needs at work. In turn, higher levels of need satisfaction should lead to lower levels of psychological distress, work-to-family conflict, and turnover intentions at the individual level, and reduce collective levels of work-to-family conflict (but not psychological distress) at the work unit level. Thus, to maximally support employees, it might be interesting to develop work cultures where employees' psychological well-being and safety are a

priority and where all levels of the organization contribute to defining the practices and procedures for the protection of employees' psychological health (Dollard & Bakker, 2010). Training could also be offered to supervisors to help them identify ways to detect and act on psychological health issues when such issues are raised (Huyghebaert et al., 2018a). Such interventions may also facilitate the emergence of a need nurturing context, in turn leading to lower risks of experiencing work-to-family conflict (Huyghebaert et al., 2018b). However, pending additional research, interventions seeking to create a need nurturing context should be done cautiously as such a work context might increase the risks of turnover. Efforts in this direction should be carefully devised to nurture need satisfaction as something that takes place within the confine of the organization.

Furthermore, from a research perspective, the adoption of a person-centered approach aiming at identifying profiles of work units characterized by different levels of job demands and resources might contribute to the identification of actionable levers of intervention to increase the likelihood of more desirable profiles. Indeed, the present research indicates that the core driver of need satisfaction is represented by an aggregated global measure of work environment quality (a global factor encompassing various demands and resources). From an organizational perspective, this observation suggests that various types of job resources might compensate for the lack of some other resources, or for the presence of specific job demands, as long as the global level of work environment quality remains satisfactory. However, this also means that it would be particularly useful for organizations to be able to ascertain their unique profile of strengths and weaknesses, in order to further improve their work environment quality. Importantly, rather than focusing on interventions designed to address each of these work environment components in an isolated manner, the ability to rely on validated intervention strategies specifically designed to capitalize on their unique profile of strengths to address their unique profiles of weaknesses would appear to be particularly useful. A person-centered approach would make it possible to identify the most common organizational configurations, and thus will be able to pave the way to the development of such targeted interventions.

Conclusion

Our findings show that employees' perceptions of their work environment, despite being anchored in well-differentiated perceptions of distinct work characteristics, seem to primarily capture a global indicator of work environment quality, which in turn seems responsible for most of the effects uncovered in this study. In turn, this gestalt of work environment perceptions seems to influence various components of employees' functioning through the mediator role of their global level of need satisfaction. However, although these effects of need satisfaction are straightforward and strictly positive at the individual level, aggregated levels of need satisfaction at the work unit level play a more complex role. On the one hand, these aggregated levels seem to create a need nurturing work context that supports employees at achieving a lower level of conflict between the demands of their work and of their family life and possibly contemplating new and more fulfilling employment opportunities. On the other hand, and possibly through forced upward social comparisons, this context also seems to increase their turnover intentions. We hope that these results will help motivate further applications of the novel analyses illustrated in this study to achieve a more comprehensive picture of the mechanisms underlying the joint JD-R/SDT framework. Importantly, managerial research is fundamentally multilevel in nature. DL-MLCFA and DL-MLSEM are effective to help researchers measure and investigate phenomena happening simultaneously at the individual and workgroup levels within organizations. We hope that, in providing an illustration of these evolving methodologies, the present study might contribute to their more widespread application.

References

Antonakis, J. (2017). On doing better: From thrill of discovery to policy implications. *The Leadership Quarterly, 28*, 5-21.

Arnaud, B., Malet, L., Teissedre, F., Izaute, M., Moustafa, F., Geneste, J., … Brousse, G. (2010). Validity study of Kessler's psychological distress scales conducted among patients admitted to French emergency department for alcohol consumption-related disorders. *Alcoholism, 34*, 1235-1245.

Asparouhov, T. (2005). Sampling weights in latent variable modeling. *Structural Equation Modeling, 12*, 411-434.

Avolio, B.J., Bass, B.M., & Jung, D.I. (1999). Re-examining the components of transformational and transactional leadership using the multifactor leadership questionnaire. *Journal of Occupational and Organizational Psychology, 72*, 441-462.

Bakker, A.B., & Demerouti, E. (2007). The Job Demands-Resources model: State of the art. *Journal of Managerial Psychology, 22*, 309-328.

Bakker, A.B., & Demerouti, E. (2017). Job demands–resources theory: Taking stock and looking forward. *Journal of Occupational Health Psychology, 22*, 273-285.

Bauer, D.J., Preacher, K.J., & Gil, K.M. (2006). Conceptualizing and testing random indirect effects and moderated mediation in multilevel models. *Psychological Methods, 11*, 142-163.

Bliese, P.D., Chan, D., & Ployhart, R. (2007). Multilevel methods: Future directions in measurement, analyses, and nonnormal outcomes. *Organizational Research Methods, 10*, 551-563.

Bosmans, M.W.G., Setti, I., Sommovigo, V., & van der Velden, P.G. (2019). Do type D personality and job demands-resources predict emotional exhaustion and work engagement? A 3-wave prospective study. *Personality and Individual Differences, 149*, 167-173.

Boudrias, J.-S., Desrumaux, P., Gaudreau, P., Nelson, K., Brunet, L., & Savoie, A. (2011). Modeling the experience of psychological health at work: The role of personal resources, social-organizational resources, and job demands. *International Journal of Stress Management, 18*, 372-395.

Boudrias, J.-S., Gaudreau, P., Desrumaux, P., Leclerc, J.-S., Ntsame-Sima, M., Savoie, A., & Brunet, L. (2014). Verification of a predictive model of psychological health at work in Canada and France. *Psychologica Belgica, 54*, 55-77.

Boudrias, V., Trépanier, S.-G., Foucreault, A., Peterson, C., & Fernet, C. (2020). Investigating the role of psychological need satisfaction as a moderator in the relationship between job demands and turnover intention among nurses. *Employee Relations, 42*, 213-231.

Caesens, G., Stinglhamber, F., & Luypaert, G. (2014). The impact of work engagement and workaholism on well-being. *Career Development International, 19*, 813-835.

Camerman, J., Cropanzano, R., & Vandenberghe, C. (2007). The benefits of justice for temporary workers. *Group & Organization Management, 32*, 176-207.

Canadian Standards Association Group & Bureau de normalisation du Québec [CSA/BNQ] (2013). *Psychological health and safety in the workplace*. Ottawa, Canada: CSA Group.

Carless, S.A., Wearing, A.J., & Mann, L. (2000). A short measure of transformational leadership. *Journal of Business & Psychology, 14*, 389-405.

Chen, X., Yuan, Y., Liu, J., Zhu, L., & Zhu, Z. (2020). Social bonding or depleting? A team-level investigation of leader self-sacrifice on team and leader work engagement. *Journal of Occupational and Organizational Psychology, 93*, 912-941.

Colarelli, S.M. (1984). Methods of communication and mediating processes in realistic job previews. *Journal of Applied Psychology, 69*, 633-642.

Colquitt, J.A. (2001). On the dimensionality of organizational justice: A construct validation of a measure. *Journal of Applied Psychology, 86*, 386-400.

Dahling, J. J., & Lauricella, T. K. (2017). Linking job design to subjective career success: A test of self-determination theory. *Journal of Career Assessment, 25*, 371-388.

Deci, E.L., Olafsen, A., & Ryan, R.M. (2017). Self-determination theory in work organizations: The state of a science. *Annual Review of Organizational Psychology and Organizational Behavior, 4*, 19-43.

De Gieter, S., Hofmans, J., & Bakker, A.B. (2018). Need satisfaction at work, job strain, and performance: A diary study. *Journal of Occupational Health Psychology, 23*, 361-372.

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

Diener, E., & Fujita, F. (1997). Social comparisons and subjective well-being. In B. P. Buunk & F. X.

Gibbons (Eds.), *Health, coping, and well-being* (pp. 329-357). Erlbaum.
Dollard, M.F., & Bakker, A.B. (2010). Psychosocial safety climate as a precursor to conductive work environments, psychological health problems, and employee engagement. *Journal of Occupational and Organizational Psychology*, *83*, 579-599.
Duriez, B., Vansteenkiste, M., Soenens, B., & De Witte, H. (2007). The social costs of extrinsic relative to intrinsic goal pursuits. *Journal of Personality*, *75*, 757-782.
Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, *44*, 350-383.
Edmondson, A., & Lei, Z. (2014). Psychological safety. *Annual Review of Organizational Psychology and Organizational Behavior*, *1*, 23-43.
Eisenberger, R., Huntington, R., Hutchinson S., & Sowa, D. (1986). Perceived organizational support. *Journal of Applied Psychology*, *71*, 500-507.
Enders, C.K. (2010). *Applied missing data analysis*. New York, NY: Guilford.
Fernet, C., Trépanier, S.-G., Austin, S., Gagné, M., & Forest, J. (2015). Transformational leadership and optimal functioning at work: On the mediating role of employees' perceived job characteristics and motivation. *Work & Stress*, *29*, 11-31.
Fujita, F. (2008). The frequency of social comparison and its relation to subjective well-being. In Eid, M., & Larsen, R.J. (Eds.), The science of subjective well-being (pp. 239-257). Guilford.
Gagné, M., Morin, A.J.S., Schabram. K., Wang, Z.N., Chemolli, E., & Briand, M. (2020). Uncovering relations between leadership perceptions and motivation under different organizational contexts: A multilevel cross-lagged analysis. *Journal of Business & Psychology*, *35*, 713-732.
Gerber, J.P., Wheeler, L., & Suls, J. (2018). A social comparison theory meta-analysis 60+ years on. *Psychological Bulletin*, *144*, 177-197.
Gillet, N., Fouquereau, E., Huyghebaert, T., & Colombat, P. (2015). The effects of job demands and organizational resources through psychological need satisfaction and thwarting. *Spanish Journal of Psychology*, *18*, 1-19.
Gillet, N., Morin, A.J.S., Choisay, F., & Fouquereau, E. (2019a). A person-centered representation of basic need satisfaction balance at work. *Journal of Personnel Psychology*, *18*, 113-128.
Gillet, N., Morin, A.J.S, Huart, I., Colombat, P., & Fouquereau, E. (2020). The forest and the trees: Investigating the globability and specificity of employees' basic need satisfaction at work. *Journal of Personality Assessment*, *102*, 702-713.
Gillet, N., Morin, A.J.S., Huyghebaert, T., Burger, L., Maillot, A., Poulin, A., & Tricard, E. (2019b). University students' need satisfaction trajectories: A growth mixture analysis. *Learning and Instruction*, *60*, 275-285.
Gonzàlez-Romà, V., & Hernàndez, A. (2017). Multilevel modeling: Research-based lessons for substantive researchers. *Annual Review of Organizational Psychology and Organizational Behavior*, *4*, 183-210.
Gonzàlez-Romà, V., & Lloret, S. (1998). Construct validity of Rizzo et al.'s (1970) role conflict and ambiguity scales: A multisample study. *Applied Psychology*, *47*, 535-545.
Grawitch, M. J., & Ballard, D. W. (Eds.). (2016). The psychologically healthy workplace: Building a win-win environment for organizations and employees. American Psychological Association.
Hancock, J.I., Allen, D.G., Bosco, F.A., McDaniel, K.R., & Pierce, C.A. (2013). Meta-analytic review of employee turnover as a predictor of firm performance. *Journal of Management*, *39*, 573-603.
Hetland, H., Hetland, J., Andreassen, C.S., Pallesen, S., & Notelaers, G. (2011). Leadership and fulfillment of the three basic psychological needs at work. *Career Development International*, *16*, 507-523.
Hooff, M.L.M., & De Pater, I.E. (2019). Daily associations between basic psychological need satisfaction and well-being at work: The moderating role of need strength. *Journal of Occupational and Organizational Psychology*. Early view. doi: 10.1111/joop.12260
Hu, L.T., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1-55.
Huynh, J.Y., Xanthopoulou, D., & Winefield, A.H. (2014). The job demands-resources model in emergency service volunteers. *Work & Stress*, *28*, 305-322.
Huyghebaert, T., Gillet, N., Lahiani, F.J., Dubois-Fleury, A., & Fouquereau, E. (2018a). Psychological safety climate as a human resource development target: Effects on workers functioning through need satisfaction and thwarting. *Advances in Developing Human Resources*, *20*, 169-181.

Huyghebaert, T., Gillet, N., Fernet, C., Lahiani, F.-J., & Fouquereau, E. (2018b). Leveraging psychosocial safety climate to prevent ill-being: The mediating role of psychological need thwarting. *Journal of Vocational Behavior, 107*, 111-125.

Huyghebaert, T., Morin, A.J.S., Forest, J., Fouquereau, E., & Gillet, N. (2022). A longitudinal examination of nurses' need satisfaction profiles: A latent transition analysis. *Current Psychology, 41*, 4837-4859.

Ivey, G.W., Blanc, J.R.S., Michaud, K., & Dobreva-Martinova, T. (2018). A measure and model of psychological health and safety in the workplace that reflects Canada's national standard. *Canadian Journal of Administrative Sciences, 35*, 509-522.

Jehn, K.A., Northcraft, G.B., & Neale, M.A. (1999). Why differences make a difference: A field study of diversity, conflict, and performance in workgroups. *Administrative Science Quarterly, 44*, 741-763.

Jungert, T., Van den Broeck, A., Schreurs, B., & Osterman, U. (2018). How colleagues can support each other's need satisfaction and motivation. *Applied Psychology, 67*, 3-29.

Kelly, A.C., Zuroff, D.C., Leybman, M.J., Martin, E.A., & Koestner, R. (2008). Satisfied groups and satisfied members: Untangling the between- and within-group effects of need satisfaction. *Journal of Applied Social Psychology, 38*, 1805-1826.

Kessler, R.C., Andrews, G., Colpe, L.J., Hiripi, E., Mroczek, D.K., Normand, S.-L.T., Walters, E.E., & Zaslavsky, A.M. (2002). Short screening scales to monitor population prevalence and trends in non-specific psychological distress. *Psychological Medicine, 32*, 959-976.

Kiersch, C.E., & Byrne, Z. (2015). Is being authentic fair? A multilevel examination of authentic leadership, justice, and outcomes. *Journal of Leadership & Organizational Studies, 22*, 292-303.

Kinnunen, U., Feldt, T., Siltaloppi, M., & Sonnentag, S. (2011). Job demands–resources model in the context of recovery: Testing recovery experiences as mediators. *European Journal of Work and Organizational Psychology, 20*, 805-832.

Klein, K.J., & Kozlowski, S.W. (2000). From micro to meso: Critical steps in conceptualizing and conducting multilevel research. *Organizational Research Methods, 3*, 211-236.

Kmieciak, R. (2022, in press). Knowledge-withholding behaviors among IT specialists: The roles of job insecurity, work overload and supervisor support. *Journal of Management & Organization*. https://doi.org/10.1017/jmo.2021.18

Knudsen, H.K., Ducharme, L.J., & Roman, P.M. (2009). Turnover intention and emotional exhaustion "at the top": Adapting the job demands-resources model to leaders of addiction treatment organizations. *Journal of Occupational Health Psychology, 14*, 84-95.

Kubicek, B., Paškvan, M., & Korunka, C. (2015). Development and validation of an instrument for assessing job demands arising from accelerated change: The intensification of job demands scale (IDS). *European Journal of Work and Organizational Psychology, 24*, 898-913.

Lapierre, L. (2013). Questionnaire sur la satisfaction professionnelle [*Professional satisfaction questionnaire*]. Ottawa, Canada: Telfer School of Business, University of Ottawa.

Li, C.-R., Lin, C.-J., Tien, Y.-H., & Chen, C.-M. (2017). A multilevel model of team cultural diversity and creativity: The role of climate for inclusion. *Journal of Creative Behavior, 51*, 163–179.

Li, P., Taris, T. W., & Peeters, M. C. W. (2020). Challenge and hindrance appraisals of job demands: One man's meat, another man's poison? *Anxiety, Stress & Coping, 33*, 31–46.

Lim, S., & Tai, K. (2014). Family incivility and job performance: A moderated mediation model of psychological distress and core self-evaluation. *Journal of Applied Psychology, 99*, 351-359.

Lüdtke, O., Marsh, H.W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B.O. (2008). The multilevel latent covariate model: A new, more reliable approach to group level effects in contextual studies. *Psychological Methods, 13*, 203-229.

Lüdtke, O., Marsh, H.W., Robitzsch, A., & Trautwein, U. (2011). A 2 X 2 taxonomy of multilevel latent contextual models: Accuracy–bias trade-offs in full and partial error correction models. *Psychological Methods, 16*, 444-467.

Lyubomirsky, S., & Ross, L. (1997). Hedonic consequences of social comparison: A contrast of happy and unhappy people. *Journal of Personality and Social Psychology, 73*, 1141-1157.

Marsh, H.W., Abduljabbar, A.S., Abu-Hilal, M.M., Morin, A.J.S., Abdelfattah, F., Leung, K.C., … Parker, P. (2013). Factorial, convergent, and discriminant validity of TIMSS math and science motivation measures. *Journal of Educational Psychology, 105*, 108-128.

Marsh, H. W., Abduljabbar, A. S., Parker, P. D., Morin, A. J. S., Abdelfattah, F., & Nagengast, B. (2014a). The Big-Fish-Little-Pond effect in mathematics: A cross-cultural comparison of U.S. and

Saudi Arabian TIMSS responses. *Journal of Cross-Cultural Psychology*, *45*, 777–804.

Marsh, H.W., Hau, K-T., & Grayson, D. (2005). Goodness of fit evaluation in structural equation modeling. In A. Maydeu-Olivares & J. McArdle (Eds.), *Contemporary psychometrics* (pp. 275-340). Erlbaum.

Marsh, H.W., Hau, K-T., & Wen, Z. (2004). In search of golden rules: Comments on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler'S (1999) findings. *Structural Equation Modeling*, *11*, 320-341.

Marsh, H.W., Kuyper, H., Morin, A.J.S., Parker, P.D., & Seaton, M. (2014b). Big-fish-little-pond social comparison and local dominance effects. *Learning and Instruction*, *33*, 50–66.

Marsh, H.W., Lüdtke, O., Robitzsch, A., Trautwein, U., Asparouhov, T., Muthén, B.O., & Nagengast, B. (2009). Doubly-latent models of school contextual effects. *Multivariate Behavioral Research*, *4*, 764-802.

Marsh, H.W., Lüdtke, O., Nagengast, B., Trautwein, U., Morin, A.J.S., Abduljabbar, A.S., & Köller, O. (2012). Classroom climate and contextual effects: Conceptual and methodological issues in the evaluation of group-level effects. *Educational Psychologist*, *47*, 106-124.

Martela, F. (2015). Fallible inquiry with ethical ends-in-view: a pragmatist philosophy of science for organizational research. *Organization Studies*, *36*, 537–563

Martins, L.L., Schilpzand, M.C., Kirkman, B.L., Ivanaj, S., & Ivanaj, V. (2012). A contingency view of the effects of cognitive diversity on team performance: The moderating roles of team psychological safety and relationship conflict. *Small Group Research*, *44*, 96-126.

McDonald, R.P. (1970). Theoretical foundations of principal factor analysis and alpha factor analysis. *British Journal of Mathematical and Statistical Psychology*, *23*, 1-21.

Morin, A.J.S., Blais, A.-R., & Chénard-Poirier, L.-A. (2022). Doubly latent multilevel procedures for organizational assessment and prediction. *Journal of Business & Psychology*, 37, 47-72.

Morin, A.J.S., Boudrias, J.-S., Marsh, H.W., McInerney, D.M., Dagenais-Desmarais, V., Madore, I., & Litalien, D. (2017). Complementary variable- and person-centered approaches to exploring the dimensionality of psychometric constructs: Application to psychological wellbeing at work. *Journal of Business and Psychology*, *32*, 395-419.

Morin, A.J.S., Marsh, H.W., Nagengast, B., & Scalas, L.F. (2014). Doubly latent multilevel analyses of classroom climate. *Journal of Experimental Education*, *82*, 143-167.

Morin, A.J.S., Myers, N.D., & Lee, S. (2020). Modern factor analytic techniques: Bifactor models, exploratory structural equation modeling (ESEM) and bifactor-ESEM. In G. Tenenbaum & R. Eklund (Eds.), *Handbook of sport psychology, 4th Ed* (pp. 1044-1073). London, UK: Wiley.

Millsap, R. (2011). *Statistical approaches to measurement invariance*. New York: Taylor & Francis.

Muthén, L., & Muthén, B. (2022). *Mplus user's* guide. Los Angeles, CA: Muthén & Muthén.

Netemeyer, R.G., Boles, J.S., & McMurrian, R. (1996). Development and validation of work-family conflict and family-work conflict scales. *Journal of Applied Psychology*, *81*, 400-410.

Nielsen, K., Nielsen, M.B., Ogbonnaya, C., Känsälä, M., Saari, E., & Isaksson, K. (2017). Workplace resources to improve both employee well-being and performance: A systematic review and meta-analysis. *Work & Stress*, *31*, 101-120.

Odle-Dusseau, H.N., Britt, T.W., & Greene-Shortridge, T.M. (2012). Organizational work–family resources as predictors of job performance and attitudes: The process of work–family conflict and enrichment. *Journal of Occupational Health Psychology*, *17*, 28-40.

Onyishi, I. E., Enwereuzor, I. K., Ogbonna, M. N., Ugwu, F. O., & Amazue, L. O. (2019). Role of career satisfaction in basic psychological needs satisfaction and career commitment of nurses in Nigeria: A self-determination theory perspective. *Journal of Nursing Scholarship*, *51*, 470-479.

O'Reilly, C., Chatman, J., & Caldwell, D. (1991). People and organizational culture: A profile comparison approach to assessing person-organization fit. *Academy of Management Journal*, *34*, 487-516.

Podsakoff, N.P., LePine, J.A., & LePine, M.A. (2007). Differential challenge stressor-hindrance stressor relationships with job attitudes, turnover intentions, turnover, and withdrawal behavior: A meta-analysis. *Journal of Applied Psychology*, *92*, 438–454.

Preacher, K.J., & Selig, J.P. (2012). Advantages of Monte Carlo confidence intervals for indirect effects. *Communication Methods and Measures*, *6*, 77-98.

Preacher, K.J., Zyphur, M., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods*, *15*, 209-233.

Quigley, N.R., Tekleab, A.G., & Tesluk, P. (2007). Comparing consensus- and aggregation-based methods of measuring team-level variables. *Organizational Research Methods, 10*, 589-608.

Reilly, M. (1982). Working wives and consumption. *Journal of Consumer Research, 8*, 407-418.

Ross, L., (1977). The intuitive psychologist and his shortcomings. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology, Vol. 10* (pp. 173-220). Academic Press.

Rupp, D.E., Shao, R., Jones, K.S., & Liao, H. (2014). The utility of a multifoci approach to the study of organizational justice: A meta-analytic investigation into the consideration of normative rules, moral accountability, bandwidth-fidelity, and social exchange. *Organizational Behavior and Human Decision Processes, 123*, 159-185.

Ryan, R.M., & Deci, E.L. (2017). *Self-Determination theory: Basic psychological needs in motivation, development, and wellness.* Guilford.

Sánchez-Cardona, I., Vera, M., Martínez-Lugo, M., Rodríguez-Montalbán, R., & Marrero-Centeno, J. (2020). When the job does not fit: The moderating role of job crafting and meaningful work in the relation between employees' perceived overqualification and job boredom. *Journal of Career Assessment, 28*, 257–276

Sánchez-Oliva, D., Morin, A.J.S., Teixeira, P.J., Carraça, E.V., Palmeira, A.L., & Silva, M.N. (2017). A bifactor exploratory structural equation modeling representation of the structure of need satisfaction at work. *Journal of Vocational Behavior, 98*, 173-187.

Schaufeli, W.B., & Taris, T.W. (2014). A critical review of the job demands-resources model: Implications for improving work and health. In G. Bauer & O. Hämmig (Eds.), *Bridging occupational, organizational and public health* (pp. 43-68). Dordrecht, NL: Springer.

Schreurs, B., van Emmerik, I.H., Van den Broeck, A., & Guenter, H. (2014). Work values and work engagement within teams. *Group Dynamics, 18*, 267-281.

Sheldon, K.M. (2005). Positive value change during college: Normative trends and individual differences. *Journal of Research in Personality, 39*, 209-223.

Sheldon, K.M., & Niemiec, C.P. (2006). It's not just the amount that counts: Balanced need satisfaction also affects well-being. *Journal of Personality and Social Psychology, 91*, 331-341.

Siemsen, E., Roth, A., & Oliveira, P. (2010). Common method bias in regression models with linear, quadratic, and interaction effects. *Organizational Research Methods, 13*, 456-476.

Simon, M., Kümmerling, A., & Hasselhorn, H.-M. (2004). Work-home conflict in the European nursing profession. *International Journal of Occupational & Environmental Health, 10*, 384-391.

Sonnentag, S. (2003). Recovery, work engagement, and proactive behavior: A new look at the interface between nonwork and work. *Journal of Applied Psychology, 88*, 518-528.

Spector, P.E. (1985). Measurement of human service staff satisfaction: Development of the Job Satisfaction Survey. *American Journal of Community Psychology, 13*, 693-713.

Thiagarajan, P., Chakrabarty, S., & Taylor, R.D. (2006). A confirmatory factor analysis of Reilly's role overload scale. *Educational and Psychological Measurement, 66*, 657-666.

Tonidandel, S., King, E.B., & Cortina, J.M. (2016). *Big data at work: The data science revolution and organizational psychology*. Routledge.

Tonidandel, S., King, E.B., & Cortina, J.M. (2018). Big data methods: Leveraging modern data analytic techniques to build organizational science. *Organizational Research Methods, 21*, 525-547.

Van den Broeck, A., Ferris, D.L., Chang, C.-H., & Rosen, C. (2016). A review of self-determination theory's basic psychological needs at work. *Journal of Management, 42*, 1195-1229.

Van den Broeck, A., Sulea, C., Elst, T.V., Fischmann, G., Iliescu, D., & De Witte, H. (2014). The mediating role of psychological needs in the relation between qualitative job insecurity and counterproductive work behavior. *The Career Development International, 19*, 526-547.

Van den Broeck, A., Vansteenkiste, M., De Witte, H., & Lens, W. (2008). Explaining the relationships between job characteristics, burnout, and engagement. *Work & Stress, 22*, 277-294

Van den Broeck, A., Vansteenkiste, M., De Witte, H., Soenens, B., & Lens, W. (2010). Capturing autonomy, competence, and relatedness at work. *Journal of Occupational and Organizational Psychology, 83*, 981-1002.

Vandenberghe, C. (1999). Organizational culture, person-culture fit, and turnover: A replication in the health care industry. *Journal of Organizational Behavior, 20*, 175-184.

Vansteenkiste, M., Neyrinck, B., Niemiec, C.P., Soenens, B., De Witte, H., & Van den Broeck, A. (2007). On the relations among work value orientations, psychological need satisfaction and job

outcomes. *Journal of Occupational and Organizational Psychology, 80*, 251-277.

Wallace, J.C., Edwards, B.D., Paul, J., Burke, M., Christian, M., & Eissa, G. (2013). Change the referent? A meta-analytic investigation of direct and referent-shift consensus models for organizational climate. *Journal of Management, 42*, 838-861.

Weale, V.P., Wells, Y.D., & Oakman, J. (2019). The work-life interface: A critical factor between work stressors and job satisfaction. *Personnel Review, 48*, 880-897.

Webber, S.S., & Donahue, L.M. (2001). Impact of highly and less- job-related diversity on work group cohesion and performance: A meta-analysis. *Journal of Management, 27*, 141-162.

Weller, S.C. (2007). Cultural consensus theory: Applications and frequently asked questions. *Field Methods, 19*, 339–368.

White, J.B., Langer, E.J., Yariv, L., & Welch, J.C. IV. (2006). Frequent social comparisons and destructive emotions and behaviors. *Journal of Adult Development, 13*, 36-44.

Xanthopoulou, D., Bakker, A.B., Demerouti, E., & Schaufeli, W.B. (2007). The role of personal resources in the job demands-resources model. *International Journal of Stress Management, 14*, 121-141.

Zhou, Z.E., Meier, L.L., & Spector, P.E. (2019). The spillover effects of coworker, supervisor, and outsider workplace incivility on work-to-family conflict. *Journal of Organizational Behavior, 40*, 1000-1012.

Zohar, D. (2000). A group-level model of safety climate: Testing the effect of group climate on microaccidents in manufacturing jobs. *Journal of Applied Psychology, 85*, 587-596.

Table 1

List of Job Demands and Resources Measured in the Present Study

Name	Type	Definition
Role ambiguity	Job Demand	The degree to which clear and consistent information is lacking about the expectations associated with one's position (Gonzàlez-Romà & Lloret, 1998).
Work overload	Job Demand	Assigning to employees a workload that is greater than the amount of time available to them to complete their work-related tasks (Kmieciak, 2022).
Psychological safety	Workgroup resource	Indicates whether it is safe for employees to take risks (e.g., speaking up, innovating) within their workgroup (Edmondson & Lei, 2014),
Interpersonal justice	Workgroup resource	A characteristic of workgroups in which employees interact respectfully with one another (Colquitt, 2001).
Contingent recognition and reward	Leader resources	Supervisor's ability to provide employees with recognition, rewards, and support for meeting expectations (Avolio et al., 1999).
Transformational leadership	Leader resource	Supervisor's ability to inspire and motivate employees' loyalty and involvement (Avolio et al., 1999).
Supervisors' safety actions and expectations	Leader resources (military)	Supervisors' actions and expectations targeted at protecting and supporting physical safety (Zohar, 2000), which also convey to employees the idea that their well-being is important (Dollard & Bakker, 2010). Important in workplaces involving risks to one's physical integrity (e.g., military).
Organizational support	Organizational resource	The extent to which the organization values and supports employees' contributions and well-being (Eisenberger et al., 1986).
Culture of respect	Organizational resource	The extent to which the organization promotes respect and interpersonal civility (O'Reilly et al., 1991). The organizational counterpart to our workgroup-level interpersonal justice resource (Rupp et al., 2014).

Table 2

Goodness-of-Fit Statistics of the Main Measurement and Predictive Models

Description	χ^2 (*df*)	CFI	TLI	RMSEA
Outcomes (Single Level)				
Confirmatory Factor Analysis (CFA)	2394.491 (130)*	.975	.971	.056
Outcomes (Multilevel)				
Multilevel CFA	3803.513 (262)*	.950	.941	.049
Invariant Loadings	3882.244 (277)*	.949	.943	.048
Need Satisfaction (Single Level)				
CFA	1351.021 (95)*	.957	.946	.048
Bifactor-CFA	666.391 (82)*	.980	.971	.035
Need Satisfaction (Multilevel)				
Multilevel CFA	2280.423 (196)*	.946	.933	.043
Invariant Loadings	2260.008 (209)*	.946	.938	.041
Multilevel Bifactor-CFA	1248.752 170)*	.972	.960	.033
Invariant Loadings	1262.797 (198)*	.972	.966	.031
Work Environment (Single Level)				
CFA	10154.834 (1178)*	.931	.925	.037
Bifactor-CFA (1 Global Factor)	8236.540 (1172)*	.945	.941	.032
Bifactor-CFA (2 Global Factors)	8135.968 (1170)*	.946	.941	.032
Bifactor-CFA (4 Global Factors)	6631.511 (1166)*	.958	.954	.029
Work Environment (Multilevel)				
Multilevel CFA	29349.690 (2357)*	.873	.862	.045
Invariant Loadings	19386.067 (2398)*	.920	.915	.035
Multilevel Bifactor-CFA (1 Global)	36722.025 (2345)*	.838	.824	.051
Invariant Loadings	16161.068 (2436)*	.935	.932	.031
Multilevel Bifactor-CFA (2 Global)	39350.928 (2343)*	.825	.810	.053
Invariant Loadings	16191.065 (2433)*	.935	.932	.031
Multilevel Bifactor-CFA (4 Global)	27762.709 (2333)*	.880	.869	.044
Invariant Loadings	13806.506 (2421)*	.946	.943	.029
Global Model				
Single Level	19288.249 (3317)*	.938	.933	.029
Multilevel (Invariant Factor Loadings)	39709.225 (6777)*	.924	.920	.029
Predictive Models (Multilevel)				
P1. Fully mediated global effects	43473.750 (6947)*	.916	.913	.030
P2. Partially mediated global effects	43952.153 (6941)*	.914	.912	.031
P3. P1 + Specific WE –> Global NS	43430.478 (6927)*	.916	.913	.030
P4. P1 + Global WE –> Specific NS	43158.491 (6941)*	.916	.913	.030
P5. P1 + Specific WE –> Specific NS	87372.817 (6887)*	.814	.807	.045
P6. P1 + Specific NS –> Outcomes	47047.639 (6929)*	.907	.904	.032
P7. P1 + Specific WE –> Outcomes	59724.728 (6887)*	.878	.873	.037

Note. * $p < .01$; χ^2: Robust chi-square test of exact fit; *df*: Degrees of freedom; CFI: Comparative fit index; TLI: Tucker-Lewis index; RMSEA: Root mean square error of approximation; CFA: Confirmatory factor analyses; WE: Work environment; NS: Need satisfaction.

Table 3

Reliability Information and Intraclass Correlations

	ω	ω_{L1}	ω_{L2}	ICC1	ICC2
WGR: Psychological Safety (S-Factor)	.550	.543	.850	.016	.645
WGR: Interpersonal Justice (S-Factor)	.874	.870	.834	.066	.890
LR: Contingent Recognition/Reward (S-Factor)	.677	.674	.803	.123	.942
LR: Transformational Leadership (S-Factor)	.918	.916	.973	.018	.674
LR: Safety Actions(S-Factor)	.641	.613	.964	.015	.634
LR: Safety Expectations (S-Factor)	.762	.760	.789	.039	.824
OR: Organizational Support (S-Factor)	.586	.578	.925	.048	.851
OR: Interpersonal Respect Culture (S-Factor)	.689	.684	.906	.073	.900
JD: Ambiguity (S-Factor)	.808	.808	.727	.022	.718
JD: Overload (S-Factor)	.904	.901	.970	.047	.850
Global Work Environment (G-Factor)	.976	.974	.994	.066	.889
Relatedness NS (S-Factor)	.809	.804	.792	.018	.672
Autonomy NS (S-Factor)	.626	.619	.870	.074	.901
Competence NS (S-Factor)	.878	.876	.943	.041	.830
Global NS (G-Factor)	.886	.881	.957	.058	.875
Psychological Distress	.930	.928	.979	.035	.804
Turnover Intentions	.825	.826	.729	.045	.843
Work-to-Family Conflict	.959	.949	.994	.188	.964

Note. G: Global factor from a bifactor measurement model; S: Specific factor from a bifactor measurement model; WGR: Workgroup resources; LR: Leader resource; OR: Organization resource; JD: Job demand; NS: Need satisfaction; ω: Omega coefficient of model-based composite reliability estimated from a single level model; ω_{L1}: Omega coefficient of model-based composite reliability estimated at level 1 (individual) from a multilevel model; ω_{L2}: Omega coefficient of model-based composite reliability estimated at level 2 (unit) from a multilevel model; ICC1: Intraclass correlation; ICC2: Reliability of unit-level aggregation.

Table 4

Multilevel Predictive Results

	Raw Multilevel Effects			Contextual Effects		
	Est. [CI]	Std. [CI]	ES [CI]	Est. [CI]	Std. [CI]	ES [CI]
Level 1 Effects						
WE—>NS	.818 [.737; .899]**	.861 [.775; .946]**	.881 [.794; .969]**			
NS—> PD	-.813 [-.910; -.716]**	-.670 [-.750; -.590]**	-.680 [-.760; -.599]**			
NS —> TI	-.900 [-.994; -.806]**	-.737 [-.814; -.660]**	-.747 [-.825; -.669]**			
NS —> WFC	-.685 [-.769; -.601]**	-.419 [-.471; -.368]**	-.468 [-.525; -.410]**			
Indirect (PD)	-.665 [-.726; -.602]**					
Indirect (TI)	-.736 [-.807; -.665]**					
Indirect (WFC)	-.561 [-.620; -.501]**					
Level 2 Effects						
Climate						
WE—>NS	.789 [.652; .927]**	.220 [.182; .259]**	.253[.209; .297]**			
Context						
NS—> PD	-.860 [-1.085; -.635]**	-.155 [-.195; -.114]**	-.157 [-.198; -.116]**	-.047 [-.242; .148]	-.008 [-.043; .027]	-.009 [-.044; .027]
NS —> TI	-.700 [-.843; -.557]**	-.125 [-.151; -.100]**	-.127 [-.153; -.101]**	.200 [.056; .344]**	.036 [.010; .062]**	.036 [.010; .062]**
NS —> WFC	-2.029 [-2.872; -1.186]**	-.271 [-.384; -.158]**	-.302 [-.428; -.177]**	-1.344 [-2.172; -.516]**	-.179 [-.290; -.069]**	-.200 [-.323; -.077]**
Indirect (PD)	-.679 [-.865; -.499]**			-.037 [-.185; .120]		
Indirect (TI)	-.552 [-.699; -.418]**			.158 [.044; .278]**		
Indirect (WFC)	-1.602 [-2.218; -.964]**			-1.061 [-1.668; -.427]**		

Note. * $p \leq .05$; ** $p \leq .01$; WE: Work environment; NS: Need satisfaction; PD: Psychological distress; TI: Turnover intentions; WFC: Work-to-family conflict; Est. = Unstandardized parameter estimate; Std. = Standardized parameter estimate; ES: Effect size estimate; CI: 95% confidence interval; confidence intervals for the coefficients themselves, as well as for the total effects, are parametric confidence intervals (i.e., estimate +/- 1.96*standard error); confidence intervals for all indirect effects are nonparametric Monte Carlo confidence intervals (Preacher & Selig, 2012).

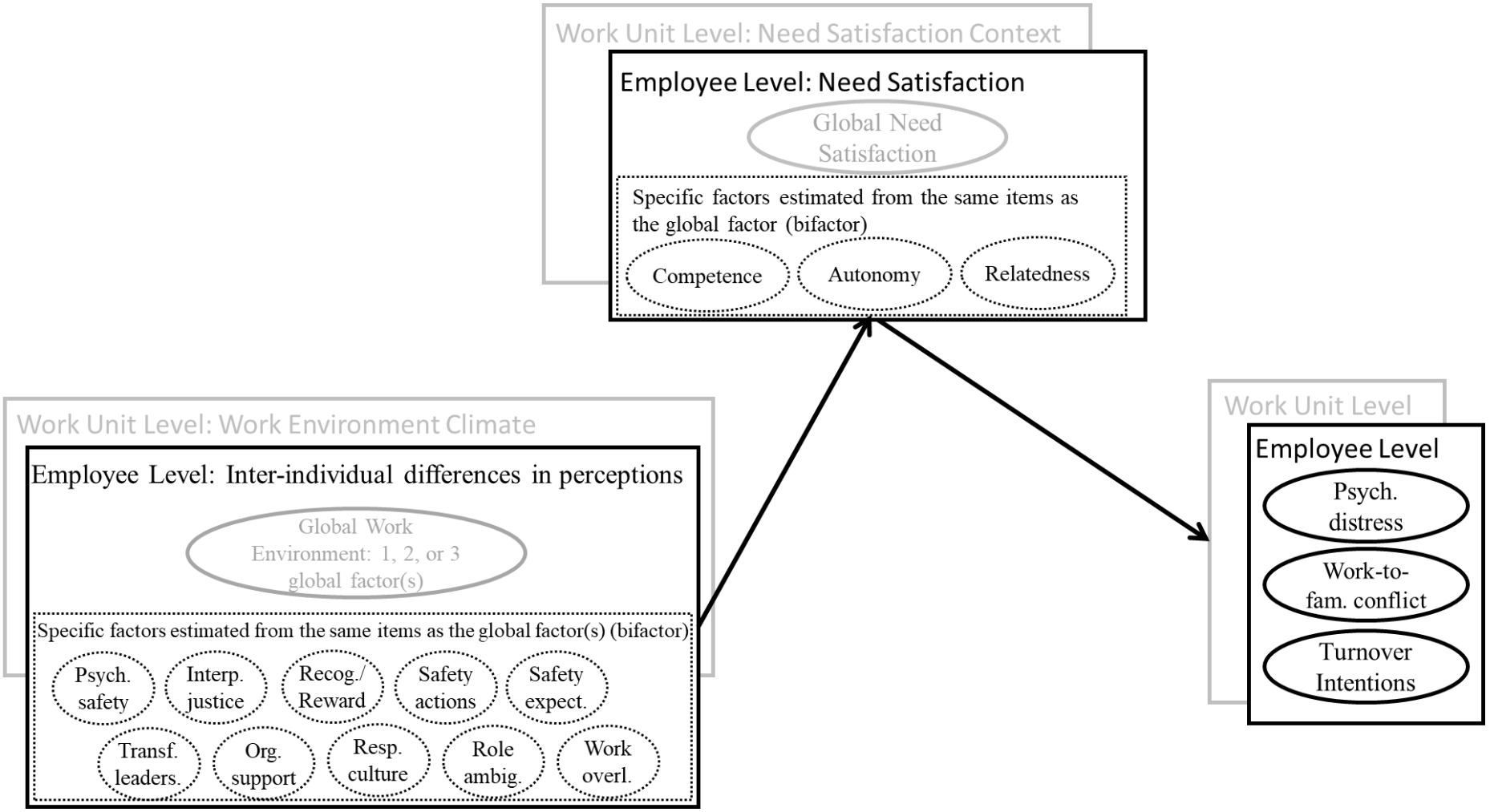

Figure 1
Theoretical Representation of the Model Considered in our Real Data Application.

Online Supplements for:

A Multilevel Perspective on the Role of Job Demands, Job Resources, and Need Satisfaction for Employees' Outcomes

Section 1
Multilevel Measurement

Conceptual Distinctiveness

It has long been considered a good practice to ascertain the distinctiveness of the measures used in a study via confirmatory factor analyses (CFA). In multilevel analyses, even when employees hold reasonably distinct perceptions of the constructs, this distinction does not necessarily translate to their aggregated perceptions, which sometimes come to reflect the global functioning of the group or organization across dimensions more than their distinctive nature (Morin et al., 2014, 2022). More reliable measurement (made possible by doubly latent methods) also tends to generate higher factor correlations, making it critical to verify the distinctiveness of multilevel constructs, something which has been rarely done in organizational research (Arthur et al., 2007; Shortell et al., 1991).

Modern factor analytic methods can be helpful when constructs that are well-differentiated conceptually overlap empirically (Morin et al., 2020). More precisely, bifactor models disaggregate ratings of multiple dimensions into non-overlapping (uncorrelated, orthogonal) global (G-factor) and specific (S-factors) components. The G-factor represents the variance shared across all items, whereas the S-factors reflect the variance shared among the items associated with a subscale left unexplained by the G-factor. These S-factors can be interpreted as deviations from the G-factor (Morin et al., 2020). Bifactor solutions should be examined whenever there are theoretical reasons for doing so, such as in research focusing on well-being (Morin et al., 2017), or psychological need satisfaction (as in the present study: Gillet et al., 2020). However, bifactor alternatives should also be considered when preliminary results reveal problematically high correlations among factors assumed to be conceptually distinct at either level of analysis, as long as their combination in a bifactor solution makes sense theoretically, or at least logically (i.e., when the G-factor can be labelled). Otherwise, researchers are left to decide between combining or deleting the overlapping variables.

Measurement Isomorphism

Isomorphism refers to constructs possessing the same measurement structure (i.e., equal factor loadings) across levels (Bliese et al., 2007; Metha & Neale, 2005). Although isomorphism is not a requirement of all doubly latent models, it has advantages. Thus, to conceptualize a construct at the individual level (L1) as a random variable with variability at the group level (L2) (i.e., to assume that the same construct is measured across levels), isomorphism is required (Metha & Neale, 2005). Furthermore, when ratings obtained at the individual level are used to obtain a picture of the group reality via a latent aggregation process, isomorphism is required to ensure comparability across levels. More importantly, the estimation procedures required to obtain a proper estimate of contextual effects (to which will turn our attention shortly) also require isomorphism. Finally, even in the absence of empirical support, imposing isomorphism has been demonstrated to result in more accurate estimates and in greater stability in the estimation process (Lüdtke et al. 2008, 2011).

References

Arthur, W., Bell, S.T., & Edwards, B.D. (2007). A longitudinal examination of the comparative criterion-related validity of additive and referent-shift consensus operationalization of team efficacy. *Organizational Research Methods, 10*, 35-58.

Bliese, P.D., Chan, D., & Ployhart, R. (2007). Multilevel methods: Future directions in measurement, analyses, and nonnormal outcomes. *Organizational Research Methods, 10*, 551-563.

Gillet, N., Morin, A.J.S, Huart, I., Colombat, P., & Fouquereau, E. (2020). The forest and the trees: Investigating the globability and specificity of employees' basic need satisfaction at work. *Journal of Personality Assessment, 102*, 702-713.

Lüdtke, O., Marsh, H.W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B.O.

(2008). The multilevel latent covariate model: A new, more reliable approach to group level effects in contextual studies. *Psychological Methods, 13*, 203-229.

Lüdtke, O., Marsh, H.W., Robitzsch, A., & Trautwein, U. (2011). A 2 X 2 taxonomy of multilevel latent contextual models: Accuracy–bias trade-offs in full and partial error correction models. *Psychological Methods, 16*, 444-467.

Metha, P.D., & Neale, M.C. (2005). People are variable too: Multilevel structural equations modeling. *Psychological Methods, 10*, 259-284.

Morin, A.J.S., Blais, A.-R., & Chénard-Poirier, L.-A. (2022). Doubly latent multilevel procedures for organizational assessment and prediction. *Journal of Business & Psychology*, 37, 47-72.

Morin, A.J.S., Boudrias, J.-S., Marsh, H.W., McInerney, D.M., Dagenais-Desmarais, V., Madore, I., & Litalien, D. (2017). Complementary variable- and person-centered approaches to exploring the dimensionality of psychometric constructs: Application to psychological wellbeing at work. *Journal of Business and Psychology, 32*, 395-419.

Morin, A.J.S., Marsh, H.W., Nagengast, B., & Scalas, L.F. (2014). Doubly latent multilevel analyses of classroom climate. *Journal of Experimental Education, 82*, 143-167.

Morin, A.J.S., Myers, N.D., & Lee, S. (2020). Modern factor analytic techniques: Bifactor models, exploratory structural equation modeling (ESEM) and bifactor-ESEM. In G. Tenenbaum & R. Eklund (Eds.), *Handbook of sport psychology, 4th Ed* (pp. 1044-1073). London, UK: Wiley.

Shortell, S.M., Rousseau, D.M., Gillies, R.R., Devers, K.J.T., & Simons, T.L. (1991). Organizational assessment in intensive care units (ICUs): Construct development, reliability, and validity of the ICU Nurse-Physician Questionnaire. *Medical Care, 29*, 706-726.

Section 2
Multilevel Reliability Estimation

When employees' ratings are used to assess constructs at the L1 and L2 levels, three types of measurement errors can influence these ratings (Lüdtke et al., 2011; Marsh et al., 2009; Morin et al., 2022). The first refers to the reliability of participants' ratings (inter-item reliability) and reflects the ratio of true score to total variance at the item level (Nunnally & Bernstein, 1994). The second source is similar, but located at the group (L2) level (Geldhof et al., 2014; Lüdtke et al., 2011). Measurement errors related to the combination of individual ratings of multiple indicators to create a L1 score are distinct from measurement errors related to the combination of the L2 aggregation of these ratings to create group-level scores. The third source stems from (dis)agreement among employees' ratings of their group (Croon & van Veldhoven, 2007; Lüdtke et al., 2008), referred to as sampling error (Gonzàlez-Romà & Hernàndez, 2017), consensus (Chan, 1998; Klein & Kozlowski, 2000), or inter-rater reliability (Gagné et al., 2020). In single-level analyses, failure to control for the first type of errors leads to downwardly biased estimates of relations. In multilevel analyses, the joint impact of these errors is more unpredictable, and may lead to spurious associations (Marsh et al., 2010). A core advantage of doubly latent analyses is linked to their ability to control for those three types of errors by using latent variables at L1 and L2 to control for inter-item reliability and a latent aggregation process allowing L2 aggregates to be controlled for inter-rater reliability (Marsh et al., 2009, 2012).

The first two sources of measurement errors, related to "inter-item" agreement, are assessed as part of traditional analyses of reliability (e.g., Cronbach alpha). Geldhof et al. (2014) recommended omega (ω; McDonald, 1970) to assess this type of measurement errors at the individual (ω_{L1}) and group (ω_{L2}) levels. Inter-item reliability (ω, ω_{L1}, ω_{L2}) can be estimated as:

$$\omega = \frac{\left(\sum_{i=1}^{k} \lambda_i\right)^2}{\left(\sum_{i=1}^{k} \lambda_i\right)^2 + \left(\sum_{i=1}^{k} \theta_{ii}\right)}$$

where λ_i reflects the standardized factor loadings associated with the ith item (out of k items) on a specific factor, and θ_{ij} are the standardized item uniquenesses, thus providing a direct estimate of the ratio of true score variance on the total variance present at the item level.

When considering inter-rater agreement, current recommendations focus on two intra-class correlation (ICC) coefficients (Marsh et al., 2012): (a) the ICC1 reflects the proportion of variance located at L2; and (b) the ICC2 reflects the reliability of the group (L2) aggregate. Inter-rater agreement (ICC1) and reliability (ICC2) can be estimated as:

$$ICC1 = \frac{\tau_x^2}{\tau_x^2 + \sigma_x^2}$$

$$ICC2 = \frac{\tau_x^2}{\tau_x^2 + \left({\sigma_x^2}/{n_j}\right)}$$

where τ_x^2 refers to the L2 variance, σ_x^2 to the L1 variance, and n_j to the average number of participants in each of the L2 units.

ICC2, ω, ω_{L1}, ω_{L2}, and can be interpreted as any other reliability estimates (e.g., alpha; i.e., ideally greater than .60 or .70) (e.g., Klein & Kozlowski, 2000; Morin et al., 2014, 2022), whereas ICC1 should ideally be greater than .10, minimally .05, to support multilevel analyses (Gonzàlez-Romà & Hernàndez, 2017; Lüdtke et al., 2008, 2011). It is important to keep in

mind that the rough interpretation guidelines typically used to make sense of these coefficients[7] are not directly relevant to doubly latent estimation. Indeed, these guidelines were proposed to support the calculation of manifest scale scores (the sum or average of items) or the combination of individual ratings to form manifest L2 aggregates (Klein & Kozlowski, 2000). These manifest aggregates or composite incorporate the measurement errors present in their components. In contrast, doubly latent multilevel confirmatory factor analyses (DL-MLCFA) and structural equation modeling (DL-MLSEM) models rely on latent measurement and latent aggregation procedures to obtain scores that are corrected for these types of errors. In plain language, doubly latent analyses remain unbiased and accurate even in the presence of low estimates of reliability (Marsh et al., 2012; Morin et al., 2014, 2022). Low reliability should even be considered as an argument in favor of doubly latent analyses, which make it possible to obtain proper estimates in spite of this unreliability. This is not to say that reliability estimates are meaningless. Far from it. On the one hand, very low reliability (e.g., .20 or .30) suggests the measures, or group members, have so little in common that their combination may not make sense. On the other hand, when doubly latent analyses prove too complex for a specific sample (an issue to which we come back in the discussion), reliability estimates should be used to guide the decision of which partial correction model to use to maximize the quality of the results (latent measurement or latent aggregation).

References

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

Croon, M.A., & van Veldhoven, M.J.P.M. (2007). Predicting group-level outcome variables from multilevel variables measured at the individual level: A latent variable multilevel model. *Psychological Methods*, *12*, 45-57.

Gagné, M., Morin, A.J.S., Schabram. K., Wang, Z.N., Chemolli, E., & Briand, M. (2020). Uncovering relations between leadership perceptions and motivation under different organizational contexts: A multilevel cross-lagged analysis. *Journal of Business & Psychology*, *35*, 713-732.

Geldhof, G.J., Preacher, K.J., & Zyphur, M.J. (2014). Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological Methods*, *19*, 72-91.

Gonzàlez-Romà, V., & Hernàndez, A. (2017). Multilevel modeling: Research-based lessons for substantive researchers. *Annual Review of Organizational Psychology & Organizational Behavior*, *4*, 183-210.

Klein, K.J., & Kozlowski, S.W. (2000). From micro to meso: Critical steps in conceptualizing and conducting multilevel research. *Organizational Research Methods*, *3*, 211-236.

Lüdtke, O., Marsh, H.W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B. (2008). The multilevel latent covariate model: A new, more reliable approach to group level effects in contextual studies. *Psychological Methods*, *13*, 203-229.

Lüdtke, O., Marsh, H.W., Robitzsch, A., & Trautwein, U. (2011). A 2 X 2 taxonomy of multilevel latent contextual models: Accuracy–bias trade-offs in full and partial error correction models. *Psychological Methods*, *16*, 444-467.

Marsh, H.W., Lüdtke, O., Nagengast, B., Trautwein, U., Morin, A.J.S., Abduljabbar, A.S., &

[7] It is generally considered that inter-item reliability should be at least .60, and ideally higher than .70, although more flexible interpretation guidelines (at least .50) have been advocated for the S-factors estimated as part of bifactor models (e.g., Morin et al., 2020; Perreira et al., 2018). Similar guidelines are generally advocated for the interpretation of ICC2 values (reflecting the inter-reliability of the L2 aggregate; e.g., Klein & Kozlowski, 2000; Morin et al., 2014, 2022). Likewise, some have argued that intraclass correlations (ICC1) values (reflecting the proportion of the variance in ratings occurring at L2) of .100 or higher (ideally) or at least .050 (minimally) are required to rely on multilevel procedures (e.g., Gonzàlez-Romà, & Hernàndez, 2017; Lütdke et al., 2008, 2011).

Köller, O. (2012). Classroom climate and contextual effects: Conceptual and methodological issues in the evaluation of group-level effects. *Educational Psychologist, 47*, 106-124.

Marsh, H.W., Lüdtke, O., Robitzsch, A., Trautwein, U., Asparouhov, T., Muthén, B.O., & Nagengast, B. (2009). Doubly-latent models of school contextual effects: Integrating multilevel and structural equation approaches to control measurement and sampling error. *Multivariate Behavioral Research, 4*, 764-802.

Marsh, H.W., Seaton, M., Kuyper, H., Dumas; F., Huguet, P., Regner, I., Buunk, A.P., Monteil, J.M., Blanton, H., & Gibbons, F.X. (2010). Phantom behavioral assimilation effects: Systematic biases in social comparison choice studies. *Journal of Personality, 78*, 671-710.

McDonald, R.P. (1970). Theoretical foundations of principal factor analysis and alpha factor analysis. *British Journal of Mathematical and Statistical Psychology, 23*, 1-21.

Morin, A.J.S., Blais, A.-R., & Chénard-Poirier, L.-A. (2022). Doubly latent multilevel procedures for organizational assessment and prediction. *Journal of Business & Psychology*, 37, 47-72.

Morin, A.J.S., Marsh, H.W., Nagengast, B., & Scalas, L.F. (2014). Doubly latent multilevel analyses of classroom climate. *Journal of Experimental Education, 82*, 143-167.

Nunnally, J.C., & Bernstein, I.H. (1994). *Psychometric theory, 3rd Ed*. New York: McGraw-Hill.

Perreira, T.A., Morin, A.J.S., Hebert, M., Gillet, N., Houle, S.A., & Berta, W. (2018). The short form of the Workplace Affective Commitment Multidimensional Questionnaire (WACMQ-S): A bifactor-ESEM approach among healthcare professionals. *Journal of Vocational Behavior, 106*, 62-83.

Section 3
Multilevel Constructs and Centering Decisions

Centering decisions, regarding how to recode scores to obtain an interpretable 0, are critical to multilevel analyses (Enders & Tofighi, 2007; Gonzàlez-Romà & Hernàndez, 2017). When using grand-mean centering, scores are rescaled by subtracting the mean of the sample, resulting in scores reflecting deviations around a sample average of 0. When using group-mean centering, scores are rescaled by subtracting the group mean (the mean of the L2 unit), resulting in scores reflecting within-group variation. When constructs are only assessed at the group level, grand-mean centering is the only option. With climate constructs, group-mean centering should be favored, as it allows L1 ratings to be expressed as deviations from aggregated group ratings, and results in L2 aggregates independent from their L1 counterparts (Enders & Tofighi, 2007; Marsh et al., 2012; Morin et al., 2014, 2022).

With contextual constructs, grand-mean centering should be favored, as it results in ratings that still contain both sources (L1 and L2) of variation (Enders & Tofighi, 2007). With this approach, the L1 ratings retain their own "identity", making it possible to obtain estimates of L2 effects that are controlled for (or partialled out from) their L1 counterparts (Enders & Tofighi, 2007; Marsh et al., 2012; Morin et al., 2014, 2022). In other words, grand-mean centering yields estimates of L2 effects that directly test for the presence of a contextual effect.

Unfortunately, via the latent aggregation process, doubly latent models automatically rely on group-mean centering. This reality, which is still not widely known, forces researchers to calculate contextual effects by subtracting the L1 effect from its L2 counterpart (Enders & Tofighi, 2007; Marsh et al., 2012), making this calculation conditional on isomorphism (Morin et al., 2014, 2022).

To our knowledge, Gonzàlez-Romà and Hernàndez (2017) are the only one who alluded to these considerations in managerial research by recommending group-mean centering, *unless one wants to test for contextual effects*. However, no clear definition of contextual effects was provided, and neither was it mentioned that doubly latent models are automatically group-mean centered.

References

Enders, C.K. & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods, 12*, 121-138.

Gonzàlez-Romà, V., & Hernàndez, A. (2017). Multilevel modeling: Research-based lessons for substantive researchers. *Annual Review of Organizational Psychology and Organizational Behavior*, *4*, 183-210.

Marsh, H.W., Lüdtke, O., Nagengast, B., Trautwein, U., Morin, A.J.S., Abduljabbar, A.S., & Köller, O. (2012). Classroom climate and contextual effects: Conceptual and methodological issues in the evaluation of group-level effects. *Educational Psychologist, 47*, 106-124.

Morin, A.J.S., Blais, A.-R., & Chénard-Poirier, L.-A. (2022). Doubly latent multilevel procedures for organizational assessment and prediction. *Journal of Business & Psychology*, *37*, 47-72.

Morin, A.J.S., Marsh, H.W., Nagengast, B., & Scalas, L.F. (2014). Doubly latent multilevel analyses of classroom climate. *Journal of Experimental Education, 82*, 143-167.

Section 4
Previous Multilevel Research on the Job Demands-Resources (JD-R) Model

Many previous attempts have been made to investigate or integrate other levels of analysis in JD-R research. For instance, when establishing the JD-R model, Demerouti et al. (2001) found similar relations between job demands and resources and employees' burnout and disengagement at the group and individual levels, showing demands to be associated with higher burnout and resources to be associated with lower disengagement. Bakker et al. (2008) found supervisor and colleague resources to be related to lower cynicism, whereas job demands were positively related to emotional exhaustion at the individual level. They also found cynicism to predict higher financial performance at the group level. Rather than focusing on global constructs reflecting job demands and resources, other studies established the multilevel role of more specific work characteristics, such as the psychological safety climate (Dollard & Bakker, 2010; Trost, 2015), time pressure and emotional demands (Kronenwett & Rigotti, 2019), autonomy support climate (Chen et al., 2018; Liu et al., 2011), leadership (Chang et al., 2013; Gagné et al., 2020; Katou et al., 2020; Kiersch & Byrne, 2015; Tuckey et al., 2012; Waldman et al., 2015), physical demands (Wöhrmann et al., 2013), cognitive demands and resources (Tuckey et al., 2012), generic job resources (Katou et al., 2021); safety actions and expectations (Bronkhorst, 2015), organizational support (El Akremi et al., 2014; Wöhrmann et al., 2013), justice (Griffin, 2010; Herr et al., 2018; Kiersch & Byrne, 2015), work overload, role ambiguity, and role conflict (Katou et al., 2021), or incivility (Griffin, 2010) in the prediction of psychological functioning or conceptually related constructs. For instance, relying on doubly latent analyses, Katou et al. (2021) reported matching effects of transformational leadership, job resources, and job demands (role conflict, role ambiguity, and work overload) on employees' engagement and burnout. This study thus supports the need for further multilevel investigations relying on a more comprehensive set of job demands and resources, coupled with an adequate measurement of climate constructs (using items referring to the L2 reality), proper centering decisions (which are not reported in their study), and a clearer demonstration of the L2 properties of their constructs (correlations and reliability).

When we consider the aforementioned research evidence, one core limitation seems to be related to the nature of the multilevel analyses conducted in these studies, which have been typically conducted while ignoring some key aspects of multilevel analyses afforded by new methodological developments. More specifically, most of the reviewed studies provided some evidence of distinctiveness at the individual level. However, very few provide evidence of conceptual distinctiveness at L2 via multilevel CFA (Gagné et al., 2020; Katou et al., 2021; Kronenwett & Rigotti, 2019), although many of them still report L2 correlations. Furthermore, in studies assessing multiple components of job demands and resources and reporting L2 correlations, many correlations were high enough (≥.600 and even ≥.800) to suggest conceptual redundancies (Chang et al., 2013; Demerouti et al., 2001; Gagné et al., 2020; Kiersch & Byrne, 2015). To our knowledge, only one previous study (Gagné et al., 2020) also reported evidence of isomorphism (i.e., the measurement models underlying the instruments used in this study were fully invariant across type of organizations). Finally, although most of the studies reviewed in the previous section provide some information related to L1 inter-item reliability and to the L1 to L2 aggregation process. However, very few reported L2 inter-item reliability (Gagné et al., 2020), or relied on latent aggregation (Gagné et al., 2020), latent measurement at L2 (Bakker et al., 2008; Demerouti et al., 2001), or doubly latent procedures (Katou et al., 2021; Kiersch & Byrne, 2015).

As noted in Section 3 of these online supplements, the distinction between climate and context, and their centering implications, is another limitation that has rarely been taken into account in previous studies. In fact, most previous studies have failed to even mention centering (Bakker et al., 2008; Demerouti et al., 2001; Dollard & Bakker, 2010; Herr et al., 2018; Katou

et al., 2021; Liu et al., 2011; Wöhrmann et al, 2013). On the positive side, most studies of the effects of job demands and resources that mentioned centering appropriately relied on group-mean centering, which is the correct approach for climate constructs (Bronkhorst, 2015; Chen et al., 2018; El Akremi et al., 2014; Gagné et al., 2020; Kiersch & Byrne, 2015; Kronenwett & Rigotti, 2019; Waldman et al., 2015). Conversely, if we considered the multilevel studies of need satisfaction described in the main manuscript, a single study adopted a grand-mean centering strategy for the study of need satisfaction (i.e., a contextual variable; Perry et al., 2018), whereas the others did not mention centering (Jungert et al., 2018) or relied on group-mean centering (Kelly et al., 2008; Schreurs et al., 2014).

In the JD-R model, job demands and resources are explicitly conceptualized as characteristics of the workgroup having an impact on the psychological functioning of group members (Chen et al., 2020). As such, as long as these characteristics are properly measured using items explicitly referring to the group reality, they form climate constructs. This climate nature is also aligned with how these characteristics have been operationalized in previous multilevel research (e.g., Bronkhorst, 2015; Chen et al., 2018; El Akremi et al., 2014; Gagné et al., 2020; Kiersch & Byrne, 2015; Kronenwett & Rigotti, 2019; Tuckey et al., 2012; Waldman et al., 2015). When properly modelled, L2 results will thus reflect the impact of perceptions shared across group members of their group reality, and their L1 counterpart will reflect inter-individual differences in these perceptions.

References

Bakker, A.B., Van Emmerik, H., & Van Riet, P. (2008). How job demands, resources, and burnout predict objective performance. *Anxiety, Stress, & Coping*, *21*, 1-16.

Bronkhorst, B. (2015). Behaving safely under pressure: The effects of job demands, resources, and safety climate on employee physical and psychosocial safety behaviors. *Journal of Safety Research*, *55*, 63-72.

Chang, W.-J.A., Wang, Y.-S., & Huang, T.-C. (2013). Work design-related antecedents of turnover intention: A multilevel approach. *Human Resource Management*, *52*, 1-26.

Chen, S.L., Shih, C.T., & Chi, N. (2018). A multilevel job demands-resources model of work engagement: Antecedents, consequences, and boundary conditions. *Human Performance*, *31*, 282-304.

Chen, X., Yuan, Y., Liu, J., Zhu, L., & Zhu, Z. (2020). Social bonding or depleting? A team-level investigation of leader self-sacrifice on team and leader work engagement. *Journal of Occupational and Organizational Psychology*, *93*, 912-941.

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

Dollard, M.F., & Bakker, A.B. (2010). Psychosocial safety climate as a precursor to conductive work environments, psychological health problems, and employee engagement. *Journal of Occupational and Organizational Psychology*, *83*, 579-599.

El Akremi, A., Colaianni, G., Portoghese, I., Galletta, M., & Battistelli, A. (2014). How organizational support impacts affective commitment and turnover among Italian nurses. *International Journal of Human Resource Management*, *25*, 1185-1207.

Gagné, M., Morin, A.J.S., Schabram. K., Wang, Z.N., Chemolli, E., & Briand, M. (2020). Uncovering relations between leadership perceptions and motivation under different organizational contexts: A multilevel cross-lagged analysis. *Journal of Business & Psychology*. doi:10.1007/s10869- 019-09649-4

Griffin, B. (2010). Multilevel relationships between organizational-level incivility, justice and intention to stay. *Work & Stress*, *24*, 309-323.

Herr, R.M., Bosch, J.A., Loerbroks, A., Genser, B., Almer, C., van Vianen, A.E.M., & Fischer, J.E. (2018). Organizational justice, justice climate, and somatic complaints: A multilevel investigation. *Journal of Psychosomatic Research*, *111*, 15-21.

Jungert, T., Van den Broeck, A., Schreurs, B., & Osterman, U. (2018). How colleagues can support each other's need satisfaction and motivation: An intervention on employee work motivation. *Applied Psychology*, *67*, 3-29.

Katou, A.A., Koupkas, M., & Triantafillidou, E. (2022). Job demands-resources model, transformational leadership and organizational performance: A multilevel study. *International Journal of Productivity and Performance Management*, *71*, 2704-2722.

Kelly, A.C., Zuroff, D.C., Leybman, M.J., Martin, E.A., & Koestner, R. (2008). Satisfied groups and satisfied members: Untangling the between- and within-group effects of need satisfaction. *Journal of Applied Social Psychology*, *38*, 1805-1826.

Kiersch, C.E., & Byrne, Z. (2015). Is being authentic fair? A multilevel examination of authentic leadership, justice, and outcomes. *Journal of Leadership & Organizational Studies*, *22*, 292-303.

Kronenwett, M., & Rigotti, T. (2019). When do you face a challenge? How unnecessary tasks block the challenging potential of time pressure and emotional demands. *Journal of Occupational Health Psychology*, *24*, 512-526.

Liu, D., Chen, X.P., & Yao, X. (2011). From autonomy to creativity: a multilevel investigation of the mediating role of harmonious passion. *Journal of Applied Psychology*, *96*, 294-309.

Perry, S.J., Rubino, C., & Hunter, E.M. (2018). Stress in remote work: Two studies testing the demand-control-person model. *European Journal of Work and Organizational Psychology*, *27*, 577-593.

Schreurs, B., van Emmerik, I.H., Van den Broeck, A., & Guenter, H. (2014). Work values and work engagement within teams. *Group Dynamics*, *18***,** 267-281.

Trost, J.K. (2015). Head over heart; heart over head: Intuition, need for cognition, psychological safety, idea generation. *Academy of Management Proceedings*, *1*, 13038.

Tuckey, M.R., Bakker, A.B., & Dollard, M.F. (2012). Empowering leaders optimize working conditions for engagement: A multilevel study. *Journal of Occupational Health Psychology*, *17*, 15-27.

Waldman, D.A., Carter, M.Z., & Hom, P.W. (2015). A multilevel investigation of leadership and turnover behavior. *Journal of Management*, *41*, 1724-1744.

Wöhrmann, A.M., Deller, J., & Wang, M. (2013). Outcome expectations and work design characteristics in post-retirement planning. *Journal of Vocational Behavior*, *83*, 219-228.

Section 5
Sample Size Considerations, Convergence, and Alternative Specifications

DL-MLCFA and DL-MLSEM are complex statistical models requiring relatively large samples, not only to achieve a sufficient level of statistical power in the estimation of the model coefficients, but also to make it possible for the models to achieve convergence. To date, the most extensive set of simulation studies on the performance of DL-MLSEM models have been conducted by Lütdke et al. (2008, 2011). In relation to sample size, their main conclusions are that: (a) many factors play a role in influencing power and convergence with doubly latent models, making it hard to come up with any simple rule of thumb regarding sample size; (b) like in any multilevel model, sample size is a function of the L1 sample size (i.e., the number of employees), of the L2 sample size (i.e., the number of work units), and of the number of L1 informants used to rate the L2 units (i.e., the number of employees in each work unit); and (c) doubly latent estimation procedures seem to perform adequately with samples including at least 50 L2 units (although 100 would be even better), each including at least 10 to 15 participants (Lüdtke et al., 2008, 2011). Since these studies, additional statistical research has been conducted on the performance of these models. These studies have generally supported Lütdke et al.'s (2008, 2011) conclusions in terms of sample size requirement to achieve accuracy, power, and convergence (Eßer et al., 2021; Holtmann et al., 2016; Hox et al., 2010; Koch et al., 2014, 2015; Li & Beretvas, 2013; Meuleman & Billiet, 2009; NcNeish, 2017; Preacher et al., 2011). Interestingly, these studies also highlighted the fact that both sample size considerations (i.e., the number of L2 units, and the number of L1 members in each L2 unit) could compensate one another, so that having more L2 units could help achieve proper estimation when fewer members are recruited within these units, just like having more members in each L2 unit could stabilize estimation when fewer of those units are available. In this study, sample size matches these recommendations. Furthermore, all of the models estimated on this study converged on proper solutions, and on reasonably small estimates of standard errors, also supporting the idea that power was sufficient.

For applications relying on smaller samples, resulting in convergence difficulties, or resulting in inflated estimates of standard errors, isomorphism has been shown to help achieve convergence and unbiased parameter estimates (Lüdtke et al., 2008, 2011). For this reason, it is possible to start with the estimation of separate DL-MLCFA solutions for different sets of constructs and then to combine these solutions (once isomorphism is established) into a single solution. This was the approach taken in the present study. In addition, doubly latent models involve two distinct types of correction for measurement errors: (a) a latent measurement process at L1 and L2 to correct for inter-item reliability; and (b) a latent aggregation process to correct for inter-rater reliability in the assessment of the L2 reality from the L1 ratings. As noted by Lüdtke et al. (2008, 2011), as well as by Marsh et al. (2009), researchers experiencing power and/or convergence issues could also rely on models including partial correction for measurement errors (see Marsh et al., 2009 for a full presentation of these models). Deciding on the type of corrections to maintain should be guided by the estimation of the three types of measurement errors presented previously, so as to correct for the most severe form of unreliability. Likewise, factor scores can be saved from preliminary single-level or multilevel CFA models to maintain some correction for measurement errors (e.g., Gagné et al., 2020).

In any case, our recommendation would be for researchers interested in the application of doubly latent procedures to start with the estimation of DL-MLCFA models to investigate the multilevel measurement structure, distinctiveness, and isomorphism of their constructs. The results from these DL-MLCFA models should be used to obtain estimates of reliability. Ideally, these models should be estimated using all constructs to be included in a study, but it is also possible to proceed from a consideration of subsets of conceptually-related constructs (like we did in the present study) for complex applications. However, by the end of these preliminary analyses, it should be possible for researchers to incorporate all constructs (once isomorphism is imposed) within the same model. Otherwise, alternative specification (manifest measurement, manifest aggregation, or factor scores) would need to be considered, ideally in a way that is guided by the results from the analyses of reliability, so as to maintain a correction for the most severe form of measurement errors. From this point forward, it should be fairly simply to convert one's final DL-MLCFA solution to one's a priori DL-MLSEM model, and to rely on the comparison of alternative predictive models.

References

Eßer, F.J., Holtmann, J., & Eid, M. (2021) A Monte Carlo simulation study on the influence of unequal group sizes on parameter estimation in multilevel confirmatory factor analysis. *Structural Equation Modeling*, *28*, 827-838.

Gagné, M., Morin, A.J.S., Schabram. K., Wang, Z.N., Chemolli, E., & Briand, M. (2020). Uncovering relations between leadership perceptions and motivation under different organizational contexts: A multilevel cross-lagged analysis. *Journal of Business & Psychology*, *35*, 713-732.

Holtmann, J., Koch, T., Lochner, K., & Eid, M. (2016). A comparison of ML, WLSMV, and Bayesian methods for multilevel structural equation models in small samples: A simulation study. *Multivariate Behavioral Research*, *51*, 661-680.

Hox, J.J., Maas, C.J.M., & Brinkhuis, M.J.S. (2010). The effect of estimation method and sample size in multilevel structural equation modeling. *Statistica Neerlandica*, *64*, 157-170.

Koch, T., Schultze, M., Burrus, J., Roberts, R.D., & Eid, M. (2015). A multilevel CFA-MTMM model for nested structurally different methods. *Journal of Educational and Behavioral Statistics*, 40, 477–510.

Koch, T., Schultze, M., Eid, M., & Geiser, C. (2014). A longitudinal multilevel CFA-MTMM model for interchangeable and structurally different methods. *Frontiers in Psychology*, *5*, 311.

Li, X., & Beretvas, N. (2013). Sample size limits for estimating upper level mediation models using multilevel SEM. *Structural Equation Modeling*, *20*, 241-264.

Lüdtke, O., Marsh, H.W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B.O. (2008). The multilevel latent covariate model: A new, more reliable approach to group level effects in contextual studies. *Psychological Methods*, *13*, 203-229.

Lüdtke, O., Marsh, H.W., Robitzsch, A., & Trautwein, U. (2011). A 2 X 2 taxonomy of multilevel latent contextual models: Accuracy–bias trade-offs in full and partial error correction models. *Psychological Methods*, *16*, 444-467.

Marsh, H.W., Lüdtke, O., Robitzsch, A., Trautwein, U., Asparouhov, T., Muthén, B.O., & Nagengast, B. (2009). Doubly-latent models of school contextual effects: Integrating multilevel and structural equation approaches to control measurement and sampling error. *Multivariate Behavioral Research*, *4*, 764-802.

McNeish, D. (2017). Multilevel mediation with small samples: A cautionary note on the multilevel structural equation modeling framework. *Structural Equation Modeling*, *24*, 609-625.

Meuleman, B., & Billiet, J. (2009). A Monte Carlo sample size study: How many countries are needed for accurate multilevel SEM? *Survey Research Methods*, *3*, 45–58.

Preacher, K.J., Zhang, Z., & Zyphur, M. (2011). Alternative methods for assessing mediation in multilevel data. *Structural Equation Modeling*, *18*, 161-182.

Table S1

Standardized Factor Loadings (λ) and Uniquenesses (δ) for the Outcomes Measurement Models

	Single Level		Multilevel: Level 1		Multilevel: Level 2	
Items	λ	δ	λ	δ	λ	δ
Psychological Distress						
Item 1	.666	.556	.657	.568	.869	.246
Item 2	.671	.549	.665	.557	.796	.366
Item 3	.726	.472	.722	.479	.944	.108
Item 4	.849	.280	.844	.287	.974	*.051*
Item 5	.680	.538	.673	.547	.812	.341
Item 6	.654	.572	.647	.582	.839	.296
Item 7	.868	.246	.864	.254	.988	*.025*
Item 8	.783	.386	.778	.395	.944	.110
Item 9	.845	.286	.842	.291	.948	.102
Item 10	.782	.388	.778	.395	.923	.149
Work-to-Family Conflict						
Item 1	.895	.199	.873	.238	.996	*.007*
Item 2	.949	.099	.937	.123	.997	*.005*
Item 3	.931	.133	.916	.161	.996	*.009*
Item 4	.904	.182	.888	.211	.991	.018
Item 5	.853	.272	.823	.323	.950	.098
Turnover Intentions						
Item 1	.665	.558	.665	.557	.588	.654
Item 2	.864	.253	.858	.264	.780	.392
Item 3	.807	.349	.816	.333	.688	.526

Note. λ: Factor loading; δ: Item uniqueness; parameter estimates that are non-statistically significant ($p > .05$) are marked in italic.

Table S2

Standardized Factor Loadings (λ) and Uniquenesses (δ) for the Need Satisfaction Bifactor Models

	Single Level			Multilevel: Level 1			Multilevel: Level 2		
Items	S-λ	G- λ	δ	S-λ	G- λ	δ	S-λ	G- λ	δ
Relatedness									
Item 1	.590	.536	.327	.582	.539	.331	.471	.838	.075
Item 2	.511	.637	.334	.502	.643	.334	.336	.828	.202
Item 3	.604	.453	.344	.597	.456	.344	.489	.718	.245
Item 4	.459	.609	.419	.452	.609	.425	.357	.924	.019
Item 5	.553	.537	.351	.551	.536	.356	.452	.846	.080
Item 6	.506	.248	.683	.499	.272	.677	.266	.279	.852
Autonomy									
Item 1	.126	.769	.392	.137	.754	.413	.204	.836	.260
Item 2	.364	.244	.780	.363	.226	.801	.494	.228	.703
Item 3	.558	.443	.476	.556	.425	.499	.841	.477	.065
Item 4	.305	.542	.613	.297	.534	.626	.472	.631	.379
Item 5	.449	.617	.417	.450	.605	.432	.632	.631	.203
Item 6	.496	.489	.479	.501	.467	.502	.767	.531	.130
Competence									
Item 1	.747	.273	.367	.747	.264	.373	.776	.337	.284
Item 2	.794	.333	.259	.793	.325	.266	.885	.447	.016
Item 3	.797	.283	.285	.796	.276	.291	.897	.383	.048
Item 4	.716	.315	.388	.716	.310	.391	.735	.391	.306

Note. G: Global factor from a bifactor measurement model; S: Specific factor from a bifactor measurement model; λ: factor loading; δ: item uniqueness; parameter estimates that are non-statistically significant ($p > .05$) are marked in italics.

Table S3

Standardized Factor Loadings (λ) and Uniquenesses (δ) for the Work Environment Bifactor Models

	Single Level			Multilevel: Level 1			Multilevel: Level 2		
Items	S-λ	G- λ	δ	S-λ	G- λ	δ	S-λ	G- λ	δ
Team Resource: Team Psychological Safety									
Item 1	.279	.599	.564	.287	.576	.585	.303	.687	.436
Item 2	.269	.611	.554	.271	.599	.568	.354	.882	*.097*
Item 3	.399	.538	.551	.393	.520	.575	.452	.676	.338
Item 4	.305	.622	.521	.316	.610	.528	.366	.796	.232
Item 5	.370	.548	.563	.361	.535	.584	.507	.847	*.026*
Item 6	.318	.441	.705	.315	.426	.719	.489	.747	.203
Item 7	.244	.701	.449	.246	.688	.466	.300	.944	*.020*
Team Resource: Interpersonal Justice									
Item 1	.641	.641	.179	.647	.627	.189	.568	.701	.187
Item 2	.687	.689	.054	.696	.677	.057	.574	.712	.164
Item 3	.636	.699	.106	.644	.690	.110	.539	.736	.168
Item 4	.461	.525	.511	.456	.509	.533	.451	.640	.387
Leader Resource: Contingent Recognition and Reward									
Item 1	.205	.764	.374	.207	.757	.384	.128	.960	*.062*
Item 2	.400	.641	.430	.404	.629	.441	.299	.952	*.004*
Item 3	.520	.488	.492	.520	.476	.503	.447	.837	*.099*
Item 4	.636	.641	.184	.642	.631	.190	.391	.788	.227
Leader Resource: Transformational Leadership									
Item 1	.525	.652	.300	.523	.646	.309	.342	.937	*.004*
Item 2	.579	.684	.197	.582	.678	.202	.351	.908	*.053*
Item 3	.576	.688	.196	.576	.686	.197	.338	.894	*.087*
Item 4	.609	.692	.151	.610	.687	.156	.371	.927	*.003*
Item 5	.576	.656	.237	.577	.650	.244	.369	.921	*.017*
Item 6	.608	.642	.218	.607	.637	.225	.392	.912	*.014*
Item 7	.602	.674	.183	.603	.669	.189	.376	.926	*.001*
Leader Resource: Supervisor's Safety Actions									
Item 1	.401	.609	.468	.375	.588	.477	.631	.659	*.044*
Item 2	.405	.577	.503	.389	.618	.503	.607	.747	.196
Item 3	.545	.420	.527	.512	.453	.533	.781	.496	.144
Item 4	.498	.141	.732	.476	.167	.746	.926	.234	*.087*
Item 5	.467	.063	.778	.447	.089	.793	.943	.135	.093
Leader Resource: Supervisor's Safety Expectations									
Item 1	.534	.399	.555	.531	.392	.564	.507	.743	*.191*
Item 2	.511	.522	.467	.518	.510	.471	.362	.707	.368
Item 3	.538	.398	.553	.536	.407	.547	.413	.623	.441
Item 4	.532	.426	.535	.530	.424	.540	.508	.805	*.094*
Item 5	.675	.476	.317	.672	.472	.326	.489	.681	.297
Organization Resource: Organizational Support									
Item 1	.453	.733	.258	.458	.719	.272	.555	.796	*.058*
Item 2	.235	.663	.505	.235	.654	.517	.333	.846	.174
Item 3	.434	.675	.356	.434	.660	.375	.569	.791	*.050*
Item 4	.328	.723	.369	.330	.709	.388	.451	.886	*.012*

	Single Level			Multilevel: Level 1			Multilevel: Level 2		
Items	S-λ	G- λ	δ	S-λ	G- λ	δ	S-λ	G- λ	δ
Organization Resource: Interpersonal Respect Culture									
Item 1	.357	.745	.317	.359	.732	.334	.369	.859	.125
Item 2	.550	.696	.213	.559	.682	.223	.575	.802	*.027*
Item 3	.481	.654	.341	.488	.636	.358	.536	.798	*.076*
Job Demand: Role Ambiguity									
Item 1	.254	-.627	.542	.255	-.619	.551	.177	-.904	*.150*
Item 2	.401	-.669	.392	.404	-.664	.396	.248	-.857	.205
Item 3	.407	-.481	.603	.421	-.467	.605	.240	-.560	.629
Item 4	.716	-.528	.208	.720	-.523	.208	.492	-.751	.194
Item 5	.713	-.596	.137	.712	-.592	.143	.434	-.759	.236
Item 6	.526	-.666	.280	.526	-.660	.288	.354	-.933	*.004*
Job Demand: Work Overload									
Item 1	.659	-.375	.425	.656	-.362	.439	.788	-.420	.203
Item 2	.876	-.167	.204	.871	-.164	.215	.982	-.179	*.004*
Item 3	.876	-.213	.187	.876	-.218	.185	.901	-.216	.142
Item 4	.758	-.357	.298	.757	-.345	.308	.853	-.376	.132
Item 5	.682	-.372	.397	.675	-.363	.412	.840	-.437	.104
Item 6	.579	-.317	.564	.580	-.305	.571	.783	-.398	.228

Note. G: Global factor from a bifactor measurement model; S: Specific factor from a bifactor measurement model; λ: factor loading; δ: item uniqueness; parameter estimates that are non-statistically significant ($p > .05$) are marked in italics.

Table S4

Latent Variable Correlations Estimated as Part of the Overall Multilevel Measurement Model (Level 1 Correlations Below the Diagonal, and Level 2 Correlations Above the Diagonal).

		1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.
1.	TR: Psychological Safety (S-Factor)		0	0	0	0	0	0	0	0	0	0
2.	TR: Interpersonal Justice (S-Factor)	0		0	0	0	0	0	0	0	0	0
3.	LR: Contingent Recognition and Reward (S-Factor)	0	0		0	0	0	0	0	0	0	0
4.	LR: Transformational Lead. (S-Factor)	0	0	0		0	0	0	0	0	0	0
5.	LR: Safety Actions(S-Factor)	0	0	0	0		0	0	0	0	0	0
6.	LR: Safety Expectations (S-Factor)	0	0	0	0	0		0	0	0	0	0
7.	OR: Organizational Support (S-Factor)	0	0	0	0	0	0		0	0	0	0
8.	OR: Interpersonal Respect (S-Factor)	0	0	0	0	0	0	0		0	0	0
9.	JD: Ambiguity (S-Factor)	0	0	0	0	0	0	0	0		0	0
10.	JD: Overload (S-Factor)	0	0	0	0	0	0	0	0	0		0
11.	Global Work Environment (G-Factor)	0	0	0	0	0	0	0	0	0	0	
12.	Relatedness NS (S-Factor)	.187	*.029*	*.058*	*.011*	*.033*	.095	*-.025*	-.142	*.022*	*-.041*	*.041*
13.	Autonomy NS (S-Factor)	-.170	-.170	.104	*-.012*	-.115	*.022*	*-.038*	-.116	*.015*	-.166	.390
14.	Competence NS (S-Factor)	*-.014*	-.047	-.142	*-.064*	*-.046*	*.024*	-.101	-.072	-.224	-.061	-.070
15.	Global NS (G-Factor)	.214	.121	-.097	*-.025*	*.025*	-.086	.101	*.040*	-.091	*.036*	.793
16.	Psychological Distress	-.085	*-.043*	*.016*	.085	.038	*.004*	-.072	*.004*	.085	.173	-.553
17.	Turnover Intentions	*-.037*	.031	*-.018*	.095	*.030*	*-.035*	-.151	*-.026*	.073	.059	-.620
18.	Work-to-Family Conflict	*-.030*	*.004*	*-.018*	.054	.073	*-.013*	*-.041*	*-.008*	*.005*	.491	-.393

Table S4 (continued)

		12.	13.	14.	15.	16.	17.	18.
1.	TR: Psychological Safety (S-Factor)	*.072*	*-.488*	*-.193*	*.574*	*-.215*	*-.290*	*-.100*
2.	TR: Interpersonal Justice (S-Factor)	*.234*	*.348*	*.160*	*-.215*	*.108*	*.318*	*.024*
3.	LR: Contingent Recognition and Reward (S-Factor)	*-.237*	*.306*	*-.609*	*-.179*	*.078*	*.278*	*.344*
4.	LR: Transformational Lead. (S-Factor)	*.280*	*-.183*	*-.108*	*.016*	*-.095*	*-.374*	*.114*
5.	LR: Safety Actions(S-Factor)	*.709*	*-.138*	*-.394*	*-.005*	*.519*	*-.101*	*.501*
6.	LR: Safety Expectations (S-Factor)	*-.257*	*-.158*	*-.125*	*-.004*	*-.193*	*.162*	*.047*
7.	OR: Organizational Support (S-Factor)	*-.388*	*-.082*	*.190*	*.370*	*-.048*	*-.665*	*-.340*
8.	OR: Interpersonal Respect (S-Factor)	*-.601*	*-.209*	*-.313*	*.365*	*-.500*	*-.086*	*.114*
9.	JD: Ambiguity (S-Factor)	*.203*	*.325*	*-.204*	*-.384*	*.005*	*.227*	*-.113*
10.	JD: Overload (S-Factor)	*-.063*	*-.092*	*-.461*	*-.001*	*.151*	*.110*	*.539*
11.	Global Work Environment (G-Factor)	*.032*	*.542*	*-.148*	.724	*-.562*	*-.482*	*-.194*
12.	Relatedness NS (S-Factor)		0	0	0	*.645*	*.061*	*.528*
13.	Autonomy NS (S-Factor)	0		0	0	*-.183*	*.033*	*-.124*
14.	Competence NS (S-Factor)	0	0		0	*-.038*	*-.178*	-.573
15.	Global NS (G-Factor)	0	0	0		-.745	-.925	*-.377*
16.	Psychological Distress	*-.041*	-.095	*-.027*	-.633		*.519*	*.520*
17.	Turnover Intentions	*-.039*	-.292	*.015*	-.600	.556		.514
18.	Work-to-Family Conflict	*.028*	-.263	*.006*	-.346	.406	.374	

Note. G: Global factor from a bifactor measurement model; S: Specific factor from a bifactor measurement model; TR: Team resources; LR: Leader resource; OR: Organization resource; JD: Job demand; NS: Need satisfaction; ω: Omega coefficient of model-based composite reliability estimated from a single level model; ω_{L1}: Omega coefficient of model-based composite reliability estimated at level 1 (individual) from a multilevel model; ω_{L2}: Omega coefficient of model-based composite reliability estimated at level 2 (unit) from a multilevel model; ICC1: Intraclass correlation; ICC2: Reliability of unit-level aggregations; parameter estimates that are non-statistically significant ($p > .05$) are marked in italics; the number 0, with no decimal, marks correlations not estimated as part of the model.

Table S5

Goodness-of-Fit Statistics of the Single Level Measurement Invariance Models

Description	χ^2 (*df*)	CFI	TLI	RMSEA	CM	$\Delta\chi^2$ (*df*)	ΔCFI	ΔTLI	ΔRMSEA
English versus French Version									
L1. Configural invariance	24996.429 (6634)*	.935	.931	.032	--	--	--	--	--
L2 Weak invariance	26154.691 (6773)*	.932	.928	.032	L1	1089.152 (139)*	-.003	-.003	.000
L3. Strong invariance	27976.843 (6839)*	.926	.922	.034	L2	1768.365 (66)*	-.006	-.006	+.002
L4. Strict invariance	28849.968 (6924)*	.923	.921	.034	L3	574.803 (85)*	-.003	-.001	.000
L5. Correlated Uniquenesses	28930.661 (6927)*	.923	.920	.034	L4	24.463 (3)*	.000	-.001	.000
L6. Latent Variance-Covariance	29209.787 (7038)*	.922	.921	.034	L5	284.956 (111)*	-.001	+.001	.000
L7. Latent Means	29530.496 (7057)*	.921	.920	.034	L6	159.038 (19)*	-.001	-.001	.000
Male versus Female Respondents									
S1. Configural invariance	22110.035 (6634)*	.934	.929	.031	--	--	--	--	--
S2 Weak invariance	22275.715 (6773)*	.934	.930	.031	S1	228.810 (139)*	.000	.001	.000
S3. Strong invariance	22840.966 (6839)*	.931	.928	.031	S2	853.656 (66)*	-.003	-.002	.000
S4. Strict invariance	22731.691 (6924)*	.932	.930	.031	S3	120.932 (85)*	+.001	+.002	.000
S5. Correlated Uniquenesses	22734.060 (6927)*	.932	.930	.031	S4	5.569 (3)	.000	.000	.000
S6. Latent Variance-Covariance	22896.895 (7038)*	.932	.931	.031	S5	187.784 (111)*	.000	+.001	.000
S7. Latent Means	23245.766 (7057)*	.931	.930	.031	S6	305.935 (19)*	-.001	-.001	.000
Military versus Civilian Respondents									
M1. Configural invariance	21238.424 (6634)*	.930	.925	.033	--	--	--	--	--
M2 Weak invariance	21315.012 (6773)*	.930	.927	.032	M1	183.274 (139)*	.000	+.002	-.001
M3. Strong invariance	21511.494 (6839)*	.930	.927	.032	M2	197.213 (66)*	.000	.000	.000
M4. Strict invariance	21731.997 (6924)*	.929	.927	.032	M3	244.582 (85)*	-.001	.000	.000
M5. Correlated Uniquenesses	21722.426 (6927)*	.929	.927	.032	M4	1.674 (3)	.000	.000	.000
M6. Latent Variance-Covariance	21885.957 (7038)*	.929	.928	.032	M5	197.926 (111)*	.000	.001	.000
M7. Latent Means	22030.070 (7057)*	.928	.927	.032	M6	129.679 (19)*	-.001	-.001	.000

Note. * $p < .01$; χ^2: Robust chi-square test of exact fit; *df*: Degrees of freedom; CFI: Comparative fit index; TLI: Tucker-Lewis index; RMSEA: Root mean square error of approximation; CM: comparison model; Δ: Change in fit relative to the CM.

Title: Sample Mplus file for Doubly Latent Multilevel Structural Equation Model.

```
DATA:
FILE IS bigdata.csv;
VARIABLE:
NAMES ARE unit qlang sex status
NEEDR1r NEEDR2 NEEDR3r NEEDR4 NEEDR5r NEEDR6
NEEDC1 NEEDC2 NEEDC3 NEEDC4
NEEDA1 NEEDA2r NEEDA3r NEEDA4 NEEDA5 NEEDA6r
RoleAm1 RoleAm2 RoleAm3 RoleAm4 RoleAm5 RoleAm6
RleOvld1 RleOvld2 RleOvld3 RleOvld4 RleOvld5 RleOvld6
IntJust1 IntJust2 IntJust3 IntJust4
psysafe1r PsySafe2 psysafe3r PsySafe4 psysafe5r PsySafe6 PsySafe7
SupAct1 SupAct2 SupAct3 SupAct4 SupAct5
supexp1r supexp2r supexp3r supexp4r supexp5r
LEADER1 LEADER2 LEADER3 LEADER4 LEADER5 LEADER6 LEADER7
POS4y POS5ry POS6y POS7ry
Reward1 reward2 reward3 reward4 Group1y Group2y Group3y
WLB1 WLB2 WLB3 WLB4 WLB5
k1 k2 k3 k4 k5 k6 k7 k8 k9 k10
TI1 TI2 TI3;
USEVARIABLES ARE
NEEDR1r NEEDR2 NEEDR3r NEEDR4 NEEDR5r NEEDR6
NEEDC1 NEEDC2 NEEDC3 NEEDC4
NEEDA1 NEEDA2r NEEDA3r NEEDA4 NEEDA5 NEEDA6r
RoleAm1 RoleAm2 RoleAm3 RoleAm4 RoleAm5 RoleAm6
RleOvld1 RleOvld2 RleOvld3 RleOvld4 RleOvld5 RleOvld6
IntJust1 IntJust2 IntJust3 IntJust4
psysafe1r PsySafe2 psysafe3r PsySafe4 psysafe5r PsySafe6 PsySafe7
SupAct1 SupAct2 SupAct3 SupAct4 SupAct5
supexp1r supexp2r supexp3r supexp4r supexp5r
LEADER1 LEADER2 LEADER3 LEADER4 LEADER5 LEADER6 LEADER7
POS4y POS5ry POS6y POS7ry
Reward1 reward2 reward3 reward4 Group1y Group2y Group3y
WLB1 WLB2 WLB3 WLB4 WLB5
k1 k2 k3 k4 k5 k6 k7 k8 k9 k10
TI1 TI2 TI3;
MISSING IS all (999);
CLUSTER = unit;
DEFINE:
STANDARDIZE NEEDR1r NEEDR2 NEEDR3r NEEDR4 NEEDR5r
NEEDR6 NEEDC1 NEEDC2 NEEDC3 NEEDC4
NEEDA1 NEEDA2r NEEDA3r NEEDA4 NEEDA5 NEEDA6r
RoleAm1 RoleAm2 RoleAm3 RoleAm4 RoleAm5 RoleAm6
RleOvld1 RleOvld2 RleOvld3 RleOvld4 RleOvld5 RleOvld6
IntJust1 IntJust2 IntJust3 IntJust4
psysafe1r PsySafe2 psysafe3r PsySafe4 psysafe5r PsySafe6 PsySafe7
SupAct1 SupAct2 SupAct3 SupAct4 SupAct5
supexp1r supexp2r supexp3r supexp4r supexp5r
LEADER1 LEADER2 LEADER3 LEADER4 LEADER5 LEADER6 LEADER7
POS4y POS5ry POS6y POS7ry
Reward1 reward2 reward3 reward4 Group1y Group2y Group3y
WLB1 WLB2 WLB3 WLB4 WLB5
k1 k2 k3 k4 k5 k6 k7 k8 k9 k10
TI1 TI2 TI3;
ANALYSIS:
ESTIMATOR = MLR;
TYPE IS TWOLEVEL;
ITERATIONS = 10000;
```

```
H1ITERATIONS = 10000;
MITERATIONS = 10000;
MODEL:
%WITHIN% !!! The within section refers to Level 1.
!!! Need satisfaction bifactor measurement model
!!! labels in parentheses are used to fix loadings to equality across levels
REL BY NEEDR2@1
NEEDR1r NEEDR3r NEEDR4 NEEDR5r NEEDR6 (l100-l104);
AUT BY NEEDA1@1
NEEDA2r NEEDA3r NEEDA4 NEEDA5 NEEDA6r (l105-l109);
COMP BY NEEDC1@1
NEEDC2 NEEDC3 NEEDC4 (l110-l112);
GNEED BY NEEDR1r@1;
GNEED BY NEEDR2* NEEDR3r NEEDR4 NEEDR5r NEEDR6  (l113-l117);
GNEED BY NEEDC1* NEEDC2 NEEDC3 NEEDC4 (l118-l121);
GNEED BY NEEDA1* NEEDA2r NEEDA3r NEEDA4 NEEDA5 NEEDA6r (l122-l127);
REL*; AUT*; COMP*; GNEED*;
METH BY NEEDR1r* NEEDR3r NEEDR5r NEEDA2r NEEDA3r NEEDA6r;
METH@1; [METH@0];
METH WITH REL@0 AUT@0 COMP@0 GNEED@0;
GNEED WITH REL@0 AUT@0 COMP@0;
REL WITH AUT@0 COMP@0;
AUT WITH COMP@0;
!!! Work environment bifactor measurement model
!!! Demands
AMB BY RoleAm1@1
RoleAm2 RoleAm3 RoleAm4 RoleAm5 RoleAm6 (al12-al16);
OVER BY RleOvld1@1
RleOvld2 RleOvld3 RleOvld4 RleOvld5 RleOvld6 (al17-al21);
AMB*; OVER*;
!!! Resources: Leader
SACT BY SupAct1@1
SupAct2 SupAct3 SupAct4 SupAct5 (al22-al25);
SEXP BY supexp1r@1
supexp2r supexp3r supexp4r supexp5r (al26-al29);
TFL BY LEADER1@1
LEADER2 LEADER3 LEADER4 LEADER5 LEADER6 LEADER7 (al30-al35) ;
SACT*; SEXP*; TFL*;
!!! Resources: Team
JUST BY IntJust1@1
IntJust2 IntJust3 IntJust4 (al36-al38);
SAFE BY psysafe1r@1
PsySafe2 psysafe3r PsySafe4 psysafe5r PsySafe6 PsySafe7 (al39-al44) ;
JUST*; SAFE*;
!!! Resources: Org
POS BY POS4y@1
POS5ry POS6y POS7ry (al45-al47);
REW BY Reward1@1
reward2 reward3 reward4 (al48-al50);
cult BY Group1y@1
Group2y Group3y (al51-al52);
pos*; rew*;cult*;
POS5ry WITH POS7ry;
!!! Global
GLOBAL BY SupAct2@1 ;
GLOBAL BY SupAct1* SupAct3 SupAct4 SupAct5 (al53-al56);
GLOBAL BY supexp1r* supexp2r supexp3r supexp4r supexp5r (al57-al61);
```

```
GLOBAL BY LEADER1* LEADER2 LEADER3 LEADER4 LEADER5 LEADER6
LEADER7 (al62-al68);
GLOBAL BY IntJust1* IntJust2 IntJust3 IntJust4 (al69-al72);
GLOBAL BY psysafe1r* PsySafe2 psysafe3r PsySafe4 (al73-al76);
GLOBAL BY psysafe5r* PsySafe6 PsySafe7 (al77-al79);
GLOBAL BY POS4y* POS5ry POS6y POS7ry (al80-al83);
GLOBAL BY Reward1* reward2 reward3 reward4 (al84-al87);
GLOBAL BY Group1y* Group2y Group3y (al88-al90);
GLOBAL BY RoleAm1* RoleAm2 RoleAm3 RoleAm4 RoleAm5 RoleAm6 (al104-al109);
GLOBAL BY RleOvld1* RleOvld2 RleOvld3 RleOvld4 RleOvld5 RleOvld6 (al110-al115) ;
GLOBAL*;
GLOBAL WITH SACT@0 SEXP@0 TFL@0 POS@0 REW@0 CULT@0 JUST@0
SAFE@0 AMB@0 OVER@0;
SACT WITH SEXP@0 TFL@0 POS@0 REW@0 CULT@0 JUST@0 SAFE@0 AMB@0
OVER@0;
SEXP WITH TFL@0 POS@0 REW@0 CULT@0 JUST@0 SAFE@0 AMB@0 OVER@0;
TFL WITH POS@0 REW@0 CULT@0 JUST@0 SAFE@0 AMB@0 OVER@0;
POS WITH REW@0 CULT@0 JUST@0 SAFE@0 AMB@0 OVER@0;
REW WITH CULT@0 JUST@0 SAFE@0 AMB@0 OVER@0;
CULT WITH JUST@0 SAFE@0 AMB@0 OVER@0;
JUST WITH SAFE@0 AMB@0 OVER@0;
SAFE WITH AMB@0 OVER@0;
AMB WITH OVER@0;
!!! Outcomes measurement model
DIST BY k1@1
k2 k3 k4 k5 k6 k7 k8 k9 k10 (l2-l10);
TI BY TI1@1
TI2 TI3 (l11-l12);
WLB BY WLB1@1
WLB2 WLB3 WLB4 WLB5 (l20-l23);
DIST*; TI*; WLB*;
k2 with k3; k5 with k6;
!!! A priori predictive paths
!!!! labels in parentheses are used later in the MODEL CONSTRAINT section to
!!!! calculate contextual, indirect, total, and standardized effects, and effects size.
dist ON gneed (ND_W);
TI ON gneed (NT_W);
WLB ON gneed (NW_W);
Gneed ON global (GN_W);
dist ON global@0 ;
TI ON global@0 ;
WLB ON global@0 ;
dist ON  SACT@0 SEXP@0 TFL@0 POS@0 REW@0 CULT@0 JUST@0;
dist ON  SAFE@0 AMB@0 OVER@0;
TI ON  SACT@0 SEXP@0 TFL@0 POS@0 REW@0 CULT@0 JUST@0;
TI ON  SAFE@0 AMB@0 OVER@0;
WLB ON  SACT@0 SEXP@0 TFL@0 POS@0 REW@0 CULT@0 JUST@0;
WLB ON  SAFE@0 AMB@0 OVER@0;
Gneed ON  SACT@0 SEXP@0 TFL@0 POS@0 REW@0 CULT@0 JUST@0;
Gneed ON  SAFE@0 AMB@0 OVER@0;
dist ON REL@0 AUT@0 COMP@0;
TI ON REL@0 AUT@0 COMP@0;
WLB ON REL@0 AUT@0 COMP@0;
REL ON SACT@0 SEXP@0 TFL@0 POS@0 REW@0 CULT@0 JUST@0;
REL ON global@0 SAFE@0 AMB@0 OVER@0;
AUT ON SACT@0 SEXP@0 TFL@0 POS@0 REW@0 CULT@0 JUST@0;
AUT ON global@0 SAFE@0 AMB@0 OVER@0;
COMP ON SACT@0 SEXP@0 TFL@0 POS@0 REW@0 CULT@0 JUST@0;
```

```
COMP ON global@0 SAFE@0 AMB@0 OVER@0;
%BETWEEN% !!! The within section refers to Level 2.
!!! Need satisfaction bifactor measurement model
!!! labels in parentheses are used to fix loadings to equality across levels
REL_B BY NEEDR2@1
NEEDR1r NEEDR3r NEEDR4 NEEDR5r NEEDR6 (l100-l104) ;
AUT_B BY NEEDA1@1
NEEDA2r NEEDA3r NEEDA4 NEEDA5 NEEDA6r (l105-l109);
COMP_B BY NEEDC1@1
NEEDC2 NEEDC3 NEEDC4 (l110-l112);
GNEED_B BY NEEDR1r@1;
GNEED_B BY NEEDR2* NEEDR3r NEEDR4 NEEDR5r NEEDR6  (l113-l117);
GNEED_B BY NEEDC1* NEEDC2 NEEDC3 NEEDC4 (l118-l121);
GNEED_B BY NEEDA1* NEEDA2r NEEDA3r NEEDA4 NEEDA5 NEEDA6r (l122-l127);
REL_B*; AUT_B*; COMP_B*; GNEED_B*;
GNEED_B WITH REL_B@0 AUT_B@0 COMP_B@0;
REL_B WITH AUT_B@0 COMP_B@0;
AUT_B WITH COMP_B@0;
!!! Work environment bifactor measurement model
!!! Demands
AMB_b BY RoleAm1@1
RoleAm2 RoleAm3 RoleAm4 RoleAm5 RoleAm6 (al12-al16);
OVER_b BY RleOvld1@1
RleOvld2 RleOvld3 RleOvld4 RleOvld5 RleOvld6 (al17-al21);
AMB_b*; OVER_b*;
!!! Resources: Leader
SACT_b BY SupAct1@1
SupAct2 SupAct3 SupAct4 SupAct5  (al22-al25);
SEXP_b BY supexp1r@1
supexp2r supexp3r supexp4r supexp5r (al26-al29);
TFL_b BY LEADER1@1
LEADER2 LEADER3 LEADER4 LEADER5 LEADER6 LEADER7  (al30-al35) ;
SACT_b*; SEXP_b*; TFL_b*;
!!! Resources: Team
JUST_b BY IntJust1@1
IntJust2 IntJust3 IntJust4 (al36-al38);
SAFE_b BY psysafe1r@1
PsySafe2 psysafe3r PsySafe4 psysafe5r PsySafe6 PsySafe7 (al39-al44) ;
JUST_b*; SAFE_b*;
!!! Resources: Org
POS_b BY POS4y@1
POS5ry POS6y POS7ry (al45-al47);
REW_B BY Reward1@1
reward2 reward3 reward4 (al48-al50);
cult_B BY Group1y@1
Group2y Group3y (al51-al52);
pos_b*; rew_b*; cult_b*;
!!! Global
GLOBAL_B BY SupAct2@1 ;
GLOBAL_B BY SupAct1* SupAct3 SupAct4 SupAct5 (al53-al56);
GLOBAL_B BY supexp1r* supexp2r supexp3r supexp4r supexp5r (al57-al61);
GLOBAL_B BY LEADER1* LEADER2 LEADER3 LEADER4 LEADER5 LEADER6
LEADER7 (al62-al68);
GLOBAL_B BY IntJust1* IntJust2 IntJust3 IntJust4 (al69-al72);
GLOBAL_B BY psysafe1r* PsySafe2 psysafe3r PsySafe4 (al73-al76);
GLOBAL_B BY psysafe5r* PsySafe6 PsySafe7 (al77-al79);
GLOBAL_B BY POS4y* POS5ry POS6y POS7ry (al80-al83);
GLOBAL_B BY Reward1* reward2 reward3 reward4 (al84-al87);
```

```
GLOBAL_B BY Group1y* Group2y Group3y (al88-al90);
GLOBAL_B BY RoleAm1* RoleAm2 RoleAm3 RoleAm4 RoleAm5 RoleAm6 (al104-
al109);
GLOBAL_B BY RleOvld1* RleOvld2 RleOvld3 RleOvld4 RleOvld5 RleOvld6 (al110-
al115) ;
GLOBAL_B*;
GLOBAL_B WITH SACT_B@0 SEXP_B@0 TFL_B@0 POS_B@0 REW_B@0
CULT_B@0 JUST_B@0 SAFE_B@0 AMB_B@0 OVER_B@0;
SACT_B WITH SEXP_B@0 TFL_B@0 POS_B@0 REW_B@0 CULT_B@0 JUST_B@0
SAFE_B@0 AMB_B@0 OVER_B@0;
SEXP_B WITH TFL_B@0 POS_B@0 REW_B@0 CULT_B@0 JUST_B@0 SAFE_B@0
AMB_B@0 OVER_B@0;
TFL_B WITH POS_B@0 REW_B@0 CULT_B@0 JUST_B@0 SAFE_B@0 AMB_B@0
OVER_B@0;
POS_B WITH REW_B@0 CULT_B@0 JUST_B@0 SAFE_B@0 AMB_B@0
OVER_B@0;
REW_B WITH CULT_B@0 JUST_B@0 SAFE_B@0 AMB_B@0 OVER_B@0;
CULT_B WITH JUST_B@0 SAFE_B@0 AMB_B@0 OVER_B@0;
JUST_B WITH SAFE_B@0 AMB_B@0 OVER_B@0;
SAFE_B WITH AMB_B@0 OVER_B@0;
AMB_B WITH OVER_B@0;
!!! Outcomes measurement model
DIST_B BY k1@1
k2 k3 k4 k5 k6 k7 k8 k9 k10 (l2-l10);
TI_B BY TI1@1
TI2 TI3 (l11-l12);
WLB_B BY WLB1@1
WLB2 WLB3 WLB4 WLB5 (l20-l23);
DIST_B*; TI_B*; WLB_B*;
!!! A priori predictive paths
!!!! labels in parentheses are used later in the MODEL CONSTRAINT section to
!!!! calculate contextual, indirect, and standardized effects, and effects size.
dist_B ON gneed_B (ND_B);
TI_B ON gneed_B (NT_B);
WLB_B ON gneed_B (NW_B);
Gneed_B ON global_B (GN_B);
dist_B ON global_B@0 ;
TI_B ON global_B@0 ;
WLB_B ON global_B@0 ;
dist_B ON  SACT_B@0 SEXP_B@0 TFL_B@0 POS_B@0 REW_B@0 CULT_B@0
JUST_B@0;
dist_B ON  SAFE_B@0 AMB_B@0 OVER_B@0;
TI_B ON  SACT_B@0 SEXP_B@0 TFL_B@0 POS_B@0 REW_B@0 CULT_B@0
JUST_B@0;
TI_B ON  SAFE_B@0 AMB_B@0 OVER_B@0;
WLB_B ON  SACT_B@0 SEXP_B@0 TFL_B@0 POS_B@0 REW_B@0 CULT_B@0;
WLB_B ON JUST_B@0 SAFE_B@0 AMB_B@0 OVER_B@0;
Gneed_B ON  SACT_B@0 SEXP_B@0 TFL_B@0 POS_B@0 REW_B@0 CULT_B@0;
Gneed_B ON JUST_B@0 SAFE_B@0 AMB_B@0 OVER_B@0;
dist_B ON REL_B@0 AUT_B@0 COMP_B@0;
TI_B ON REL_B@0 AUT_B@0 COMP_B@0;
WLB_B ON REL_B@0 AUT_B@0 COMP_B@0;
REL_B ON SACT_B@0 SEXP_B@0 TFL_B@0 POS_B@0 REW_B@0 CULT_B@0
JUST_B@0;
REL_B ON global_B@0 SAFE_B@0 AMB_B@0 OVER_B@0;
AUT_B ON SACT_B@0 SEXP_B@0 TFL_B@0 POS_B@0 REW_B@0 CULT_B@0
JUST_B@0;
AUT_B ON global_B@0 SAFE_B@0 AMB_B@0 OVER_B@0;
```

```
COMP_B ON SACT_B@0 SEXP_B@0 TFL_B@0 POS_B@0 REW_B@0 CULT_B@0;
COMP_B ON JUST_B@0 global_B@0 SAFE_B@0 AMB_B@0 OVER_B@0;
MODEL CONSTRAINT:
!!!! Calculation of contextual effects.
NEW (ND_C NT_C NW_C);
ND_C = ND_B - ND_W;
NT_C = NT_B - NT_W;
NW_C = NW_B - NW_W;
!!!! Calculation of indirect and total effects based on L1, L2 and contextual effects.
NEW (ind_DW ind_DB ind_DC ind_DT1);
ind_DW = ND_W * GN_W;
ind_DB = ND_B * GN_B;
ind_DC = ND_C * GN_B;
ind_DT1 = ind_DW + ind_DC;
NEW (ind_TW ind_TB ind_TC ind_TT1);
ind_TW = NT_W * GN_W;
ind_TB = NT_B * GN_B;
ind_TC = NT_C * GN_B;
ind_TT1 = ind_TW + ind_TC;
NEW (ind_WW ind_WB ind_WC ind_WT1);
ind_WW = NW_W * GN_W;
ind_WB = NW_B * GN_B;
ind_WC = NW_C * GN_B;
ind_WT1 = ind_WW + ind_WC;
!!! Calculation of L1 standardized effects and effect sizes.
!!! Variance estimates taken from the multilevel CFA model.
NEW (S_NDW ES_NDW S_NTW ES_NTW S_NWW ES_NWW S_GNW ES_GNW);
S_NDW = ND_W * (sqrt(.294)/sqrt(0.433));
ES_NDW = ND_W * (sqrt(.294)/sqrt(0.421));
S_NTW = NT_W * (sqrt(.294)/sqrt(0.438));
ES_NTW = NT_W * (sqrt(.294)/sqrt(0.427));
S_NWW = NW_W * (sqrt(.294)/sqrt(0.785));
ES_NWW = NW_W * (sqrt(.294)/sqrt(0.631));
S_GNW = GN_W * (sqrt(.341)/sqrt(0.308));
ES_GNW= GN_W * (sqrt(.341)/sqrt(0.294));
NEW (S_NDB ES_NDB S_NTB ES_NTB S_NWB ES_NWB S_GNB ES_GNB);
S_NDB = ND_B * (sqrt(.014)/sqrt(0.433));
ES_NDB = ND_B * (sqrt(.014)/sqrt(0.421));
S_NTB = NT_B * (sqrt(.014)/sqrt(0.438));
ES_NTB = NT_B * (sqrt(.014)/sqrt(0.427));
S_NWB = NW_B * (sqrt(.014)/sqrt(0.785));
ES_NWB = NW_B * (sqrt(.014)/sqrt(0.631));
S_GNB = GN_B * (sqrt(.024)/sqrt(0.308));
ES_GNB= GN_B * (sqrt(.024)/sqrt(0.234));
NEW (S_NDC ES_NDC S_NTC ES_NTC S_NWC ES_NWC);
S_NDC = ND_C * (sqrt(.014)/sqrt(0.433));
ES_NDC = ND_C * (sqrt(.014)/sqrt(0.421));
S_NTC = NT_C * (sqrt(.014)/sqrt(0.438));
ES_NTC = NT_C * (sqrt(.014)/sqrt(0.427));
S_NWC = NW_C * (sqrt(.014)/sqrt(0.785));
ES_NWC = NW_C * (sqrt(.014)/sqrt(0.631));OUTPUT:
SAMPSTAT STANDARDIZED RESIDUAL CINTERVAL;
MODINDICES (3.0) TECH1 TECH2 TECH3 TECH4 SVALUES;
```

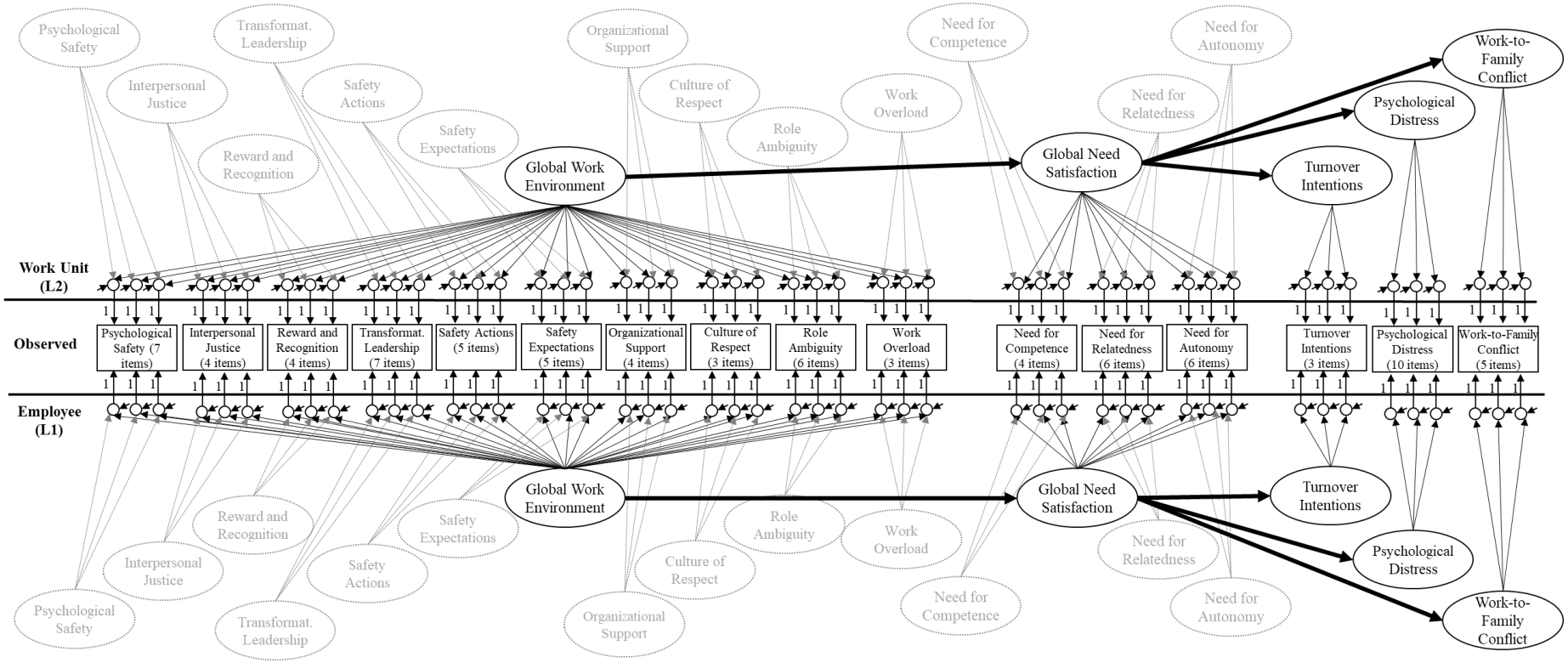

Figure S1. Doubly latent multilevel data analytic model.

Note. Item responses for each subscale are indicated in the rectangle appearing in the middle of the figure, and disaggregated across levels using a latent aggregation process (the small circles reflect the Level 1 and Level 2 components of these ratings); Latent factors are represented in ovals; Factor loadings linking items to the factors are represented as small black arrows (the scale of all factors was set by fixing the factor loading of one referent indicator to one); The predictive paths tested in the present study are represented by bold arrows; small black arrows connected to a single oval or circle reflect residual variances; to maximize the readability of the figures, factor loadings involving only three item per subscales are shown, and factors (ovals, names, and loadings) associated with the specific factors (S-factors) from bifactor measurement models are illustrated in greyscale with small dotted arrows.