Journal of Management
Vol. 47 No. 6, July 2021 1630–1656
DOI: 10.1177/0149206320916767

Incorporating Job Demand Variability Into Job Demands Theory: A Meta-Analysis

Patrick E. Downes
Texas Christian University
Cody J. Reeves
Brigham Young University
Brian W. McCormick
Northern Illinois University
Wendy R. Boswell
Texas A&M University
Marcus M. Butts
Southern Methodist University

A rich history of research on job demands suggests that employees' demands at work are related to their strain and engagement. This research often considers job demands to be fixed and stable over the course of workers' experiences, despite the existing research showing that some employees experience high levels of job demands one day and low levels the next. We seek to extend research on job demands by introducing the idea that different employees experience different levels of job demand variability (i.e., variability in job demands over workers' daily experiences). Relying on arousal theory, we posit that job demand variability moderates the between-person effects of overall job demand levels on employee strain and engagement. To test our theory, we conduct a meta-analytic path analysis of the existing experience sampling methodology research on challenge and hindrance job demands. Results show that the between-person effects of challenge (on strain and engagement) and hindrance demands (on engagement only) are stronger in studies where those demands have higher levels of daily within-person variability. Unexpectedly, the relationship between hindrance demands and strain was similarly strong across lower and higher degrees of variability. Our study suggests a need for more nuanced theory that explains how job demand variability plays a role in employee outcomes.

Supplemental material for this article is available with the manuscript on the JOM website.

Corresponding author: Patrick E. Downes, Texas Christian University, TCU Box 298530, Fort Worth, TX 76129, USA.

E-mail: patrick.downes@tcu.edu

Further, we conducted a simulation study to validate our methodology, offering utility to the broader management literature applying meta-analysis to study within-person variability. We discuss theoretical and practical implications as well as several directions for theory in this new line of reasoning.

Keywords: *motivation; meta-analysis; well-being*

Scores of academic articles spanning decades of research reinforce the idea that employees are affected by demands they experience at work. Job demands have been linked to key organizational outcomes including employee engagement and burnout (Crawford, LePine, & Rich, 2010), performance (LePine, Podsakoff, & LePine, 2005), citizenship and counterproductive work behaviors (Rodell & Judge, 2009), and workplace safety (Nahrgang, Morgeson, & Hofmann, 2011). The job demands perspective has taken a prominent place in management research (see Parker, Morgeson, & Johns, 2017), and a key understanding is that different types of job demands produce distinct outcomes for individuals. Demands that promote growth and achievement (challenge demands) produce different effects than do those produced by demands that inhibit growth and achievement (hindrance demands) (Podsakoff, LePine, & LePine, 2007). This avenue of research has articulated two descriptive attributes of job demands: the level of the job demand (as some individuals have more frequent or more difficult job demands) and the type of job demand (i.e., challenge vs. hindrance demands). According to this framework, the combination of the level and type of job demand can be used to understand individuals' motivation (Bakker & Demerouti, 2007; Crawford et al., 2010).

We broaden this understanding of individuals' job demands by highlighting that job demands can vary over the course of individuals' work experiences and showing that this variability plays a role in how demands affect workers. A work environment where every day is similarly difficult as the next is much different than one where demands are inconsistent throughout a given week, even if the overall level of job demands is identical. As an example, consider that the emergency room and the operating room are both extremely demanding environments calling for precise coordination and skilled work. Yet employees in one room are accustomed to relatively regular, consistent levels of demands whereas employees in the other may have performance episodes of extremely high levels of demands followed by periods of very low demands. A key question stemming from this logic involves how these two different experiences—one more steady and even, and the other dramatically dynamic and variable—affect employees' motivation. Even if two environments have the same overall average levels of demands, the variability in those demands may result in an entirely different experience.

The broader theoretical principle in this line of reasoning is that the presence of a certain level of job demands at a given time does not have a single objective meaning but rather is interpreted relative to the moments that surround it. The waiter's lunch rush would not be a rush if the restaurant was always busy, and working a 12-hour day is an entirely different experience for people who usually work 8 hours per day than for people who usually work 12 hours per day. This is reflected in a growing within-person perspective emphasizing individuals' varying experiences over time and the way in which that variability affects outcomes (Dalal, Bhave, & Fiset, 2014). Moreover, job demand variability may be particularly important in the modern context of work where jobs, and employees' experiences in them, are changing more rapidly than ever. Alongside the growth of freelancing, flexible work practices, and gig economy roles (Bidwell, 2013; Cappelli & Keller, 2013), workers are encountering a faster pace of work as well as a faster rate of change at work. As individuals encounter more variable job demands, theory must be broadened to better understand the role that dynamism plays in workers' experiences.

Despite the intuitive appeal of job demand variability as a component of job demands, existing research relies on either one of two perspectives (i.e., within- vs. between-person differences) when a more integrated perspective is needed. On the one hand, in the between-person stream of research, job demands theory contends that some individuals have a high level of job demands and others have lower levels (e.g., Demerouti, Bakker, Nachreiner, & Schaufeli, 2001), leading to between-individual variation in individual outcomes (Crawford et al., 2010; LePine et al., 2005; Podsakoff et al., 2007). In this way, the between-persons perspective focuses solely on average levels of job demands and outcomes. On the other hand, the within-person stream of research recognizes and theorizes about within-person variation in demands (e.g., Bakker & Sanz-Vergel, 2013; Rodell & Judge, 2009), yet that research focuses primarily on the variable effects of daily job demands without fully considering how day-to-day fluctuations in job demands manifest between individuals. In this way, the within-person perspective focuses solely on deviations from individuals' average levels of demands. Although both streams shed light on job demands, they consider only one level in isolation and thus do not acknowledge the possibility that people experience variable job demands differently than they experience stable job demands (i.e., that the deviations give a new meaning to individuals' averages).

To extend this literature, we offer an integrated within- *and* between-person perspective of job demands. This new perspective offers three contributions. First, by integrating these perspectives via meta-analyses at both the within- and between-levels, we broaden extant conceptualizing, offering new theory and practical insight about how fluctuations in job demands—and not solely the average level of the demands themselves—play a role in the employee outcomes of strain and engagement. Given the centrality of job demands research to the field of management (see Parker et al., 2017—Demerouti, Bakker, Nachreiner, & Schaufeli [2001] is among the most cited articles in management), articulating and examining this novel boundary condition offers a meaningful contribution in moving forward a key area of management research. Second, this approach represents a novel way to integrate the between- and within-person perspectives that can be further applied in other areas of management research. There are many organizational phenomena that are known to vary over time (Dalal et al., 2014; McCormick, Reeves, Downes, Li, & Ilies, 2020) and could be viewed through the perspective we propose. For example, leadership and social interactions are known to have dynamic elements. Viewing these phenomena through an integrated within- and between-person lens would explore how the degree of dynamism shapes these social experiences. Although not the focus of our research, we offer a template for a novel way to view management phenomena and extend theory by integrating the within- and between-person perspectives. Third, our research makes a methodological contribution consistent with this proposed perspective. We conduct a meta-analysis of experience sampling methodology (ESM) studies, and we validate our technique (through a supplemental simulation study) for future researchers' use.

Theoretical Background

Research suggests that life's demands lead to a variety of emotional and cognitive responses that tax the resources available to individuals (Lazarus & Folkman, 1984). At work, organizational researchers have examined job demands, which are aspects of a job that require sustained effort and thus produce psychological strain (Demerouti et al., 2001). Job demands have been further categorized within the challenge-hindrance demands framework (Cavanaugh, Boswell, Roehling, & Boudreau, 2000). Challenge demands are those that can promote mastery, personal growth, or future gains (Crawford et al., 2010). Hindrance demands, in contrast, are demands that have a threatening potential that can inhibit personal growth, learning, and work goal attainment.[1] Differentiating challenge demands from hindrance demands has been useful for clarifying how different types and levels of job demands are related to different individual outcomes (cf. Boswell, Olson-Buchanan, & LePine, 2004; Cavanaugh et al., 2000).

The challenge-hindrance framework primarily emphasizes strain and employee engagement as outcomes of job demands. Strain represents individuals' psychological or physiological reaction to an experienced stress such as a job demand (cf. Scheck, Kinicki, & Davy, 1995; Spector, Chen, & O'Connell, 2000). Job engagement represents employees' harnessing of their whole selves to their work roles whereby they fully apply their physical, cognitive, and emotional resources toward their role performance (Kahn, 1990; Rich, Lepine, & Crawford, 2010). Most notably, both challenge and hindrance demands are theorized to positively relate to strain because both types of demands tax individuals and consume personal resources (Boswell et al., 2004). However, challenge demands are theorized to positively relate to job engagement, whereas hindrance demands should negatively relate to job engagement. This is because challenge demands, although taxing personal resources, have "associated personal gains" (Cavanaugh et al., 2000: 68), which lead to positive reactions and a more active style of coping (Crawford et al., 2010). In contrast, engagement should weaken as individuals detach or redirect energy when faced with hindrance demands (Crawford et al., 2010). Meta-analyses have confirmed the core elements of this challenge-hindrance framework (Lepine et al., 2005; Podsakoff et al., 2007); challenge demands, though they produce strain, can be beneficial for engagement, whereas hindrance demands produce strain and inhibit engagement.

Meta-analyses in this area (i.e., Crawford et al., 2010; LePine et al., 2005; Podsakoff et al., 2007) focus on between-person differences in job demands. In this perspective, job demands lead to between-individual variation in outcomes (e.g., people who face greater challenge demands generally experience greater strain and engagement). By virtue of their empirical analyses of one-time surveys, this literature has generally assumed demands are stable. However, the within-person perspective has shown day-to-day changes in job demands (e.g., Kühnel, Sonnentag, & Bledow, 2012; Rodell & Judge, 2009). This perspective articulates that job demands vary over time yet focuses on how the momentary level of job demands at one time predicts the momentary level of outcomes at the same (or another) point in time. This improves upon cross-sectional research by offering a more ecologically valid test (see Beal, 2015) of the challenge-hindrance framework, but it *accounts* for variability in job demands rather than substantively exploring how variability influences the job demand experience.

Considering Job Demand Variability in Evaluating the Effects of Job Demand Level

Job demand variability occurs as workers' job demands fluctuate across performance episodes. For some individuals, these swings in demands can be quite drastic, with some performance episodes having high job demands and others having low job demands. For bartenders, grocers, or retail salespeople, some days will be busy and others slow. Even beyond workload demands, employees can experience some periods with high levels of office politics or relational conflict (perhaps surrounding scheduled events or organizational initiatives) and others free from such hassles. For others, job demands are stable and constant where workers spend every day working through the same checklist with the same individuals and the same workplace dynamics they experienced the day before.

We propose conceptualizing this variability as a component of the job demands experience, differentiating between the level of job demands and the variability in those job demands over time. At present, definitions of job demands (e.g., Crawford et al., 2010; Podsakoff et al., 2007) emphasize between-job (and at the same time between-person) demand levels, which disregards any potential variation in those demands. Several areas of management research (such as team composition, which distinguishes between the overall level and diversity in member attributes [e.g., Barrick, Stewart, Neubert, & Mount, 1998] or compensation, which distinguishes between the overall level and dispersion in pay [e.g., Bloom, 1999]) have been advanced by conceptually separating the overall level of a phenomena from the variability across its lower level units. Applying similar logic to job demands, our overarching question involves the extent to which variable versus stable job demands have different effects.

Job Demand Variability and Arousal Theory

Our contention is that dealing with within-person job demand variability presents an additional imposition on personal resources that should magnify the effects of those job demands (Bargh & Chartrand, 1999). In contrast, stable job demands should promote consistency, predictability, and adaptation and should thus have weaker effects on strain and engagement. We base this proposition on arousal theory (Berlyne, 1960), which deals directly with how individuals attend to environmental stimuli that are changing over time. Arousal theory specifically posits that stimuli that deviate from a baseline command more attention and consume more personal resources (e.g., cognitive, emotional, or physical) than those that are more consistent with the baseline. One explanation for this effect lies in the dual processes involved in cognition. Specifically, individuals rely on an automatic form of processing for repeated or routine stimuli, whereas individuals often must engage a deliberate, rational system as they encounter novel stimuli (Schneider & Shiffrin, 1977). Any environmental change creates uncertainty surrounding the match between task demands and personal resources, necessitating a reappraisal of the personal resources that will be necessary to overcome the demand. Managing this reappraisal process presents an additional imposition on personal resources when stimuli are variable because they present ambiguity that would not be present were environmental demands consistent over time. Indeed, research confirms that novel stimuli elicit more intense cognitive, affective, and physiological reactions than do repeated stimuli (Berlyne, Borsa, Craw, Gelman, & Mandell, 1965; Berlyne, Borsa, Hamacher, & Koenig, 1966; Berlyne & Carey, 1968).

Arousal theory depicts the novelty of a stimulus in three ways that rely both upon individuals' past encounters with a stimulus as well as the degree of change in stimuli experienced over time. Berlyne (1960) refers to stimuli that individuals first encounter as having "complete novelty" because individuals are completely unfamiliar with the stimuli until they are encountered. Following a first encounter, Berlyne indicates that stimuli can still have short- or long-term novelty when contrasted to the typical baseline. Short-term novelty involves stimuli changing over periods of moments (e.g., throughout an experimental session or single work episode), whereas long-term novelty describes stimuli that change between work episodes (e.g. day-to-day fluctuation). In either short- or long-term novelty, individuals are familiar with the stimulus because they have seen it before. Yet the stimulus is considered novel because it deviates from other stimuli either earlier within the work episode (short-term novelty) or in some prior work episode (long-term novelty). Importantly, arousal theory argues that stimuli with any type of novelty (whether complete, short- or long-term) have a greater potential for arousing individuals' personal resources than do stimuli that are constant.

Interestingly, research suggests that novelty impacts arousal regardless of whether individuals expect a forthcoming change. For instance, Helton, Shaw, Warm, Matthews, and Hancock (2008) experimentally manipulated the expectation of a workload change and found a nonsignificant effect of the expectedness of the change on participant stress. This finding is consistent with Berlyne's (1960) prediction that even if individuals are aware of an upcoming change in demands, the "anticipatory" arousal still burdens personal resources, thus mitigating any potential advantage of having foreknowledge that demands will change in the future (Berlyne, 1960). Therefore, the research predicts that even if people are accustomed to job demand variability in their work experiences, they are still likely to experience variable demands more strongly than people experience stable demands.

Applying arousal theory to job demands raises a question that has yet to be asked within the challenge-hindrance framework. Given the ample research suggesting that job demands vary over time (e.g., Beehr, Jex, Stacy, & Muray, 2000; Cox-Fuenzalida, 2007; Hauck, Snyder, & Cox-Fuenzalida, 2008; Shaw & Weekely, 1985), it would be valuable to know whether varying job demands exert different effects on employees than do stable job demands. Our core proposition, rooted in arousal theory, is that more variable job demands should have different effects than more stable job demands. By integrating extant theory on job demands (Crawford et al., 2010) with arousal theory (Berlyne, 1960), we form predictions below about how job demand variability moderates the effects of job demands on strain and engagement.

Challenge demands produce engagement because they promote personal growth and achievement (Crawford et al., 2010; Podsakoff et al., 2007). Yet challenge demands that represent a more stable aspect of individuals' work environments should not promote growth and achievement and should thus not stimulate increased engagement. For this reason, more variable challenge demands, which deviate more from employees' usual experiences, should have stronger effects on engagement than more stable challenge demands to which employees are more accustomed. Similarly, hindrance demands that vary over time inhibit engagement because they serve as novel or unpredictable interfering obstacles beyond employees' typical experiences. Even in a situation when hindrance demands are lower than usual, the deviation from typicality could present uncertainty, engaging deliberative cognitive processes that consume greater personal cognitive resources (Miceli & Castelfranchi, 2005). Conversely, hindrance demands that are recurrent and ongoing often are characteristic of the work environment (i.e., "the way things are around here"), and these more stable hindrance demands that characterize a work environment should not draw workers' attention (Bargh & Chartrand, 1999). Because relatively stable hindrance demands do less to arouse individuals cognitively, affectively, or physically (relative to variable demands), they should have weaker effects on engagement. We thus expect the positive effects of challenge demands and the negative effects of hindrance demands on employee engagement to be stronger when they diverge more from employees' typical experiences (i.e., high job demand variability) but weaker when they are more consistent with employees' characteristic experiences (i.e., low job demand variability).

We also expect both types of job demands to have stronger effects on strain when job demands vary more over time. From the arousal theory perspective, job demands with higher levels of variability command more attention and consume more cognitive resources than do consistent or stable demands. Building on this perspective, Helton et al. (2008) argued that variable or unpredictable demands complicate individuals' assessments of whether their personal resources are adequate for task demands. As individuals devote additional resources to assessing the match between personal resources and task demands—as would be expected in a situation where demands vary—the burden placed on cognitive resources grows, taxing individuals further and elevating their strain. Researchers have found that this mismatch occurs both when demands are higher than the baseline and when demands are lower than the baseline (cf. Miceli & Castelfranchi, 2005). This additional consumption of cognitive resources should amplify the effect of job demands on strain when the demands are variable rather than stable.

These arguments lead us to expect that people with higher levels of challenge demands should experience more strain and more engagement (relative to people with lower levels of challenge demands) primarily when that level of challenge demands varies over time. Similarly, we contend that the level of hindrance demands individuals experience should lead to strain and diminished engagement primarily when those demands vary. When challenge and hindrance demands are more stable across people's experiences, we expect the effects on strain and engagement to be weaker than when they vary over time.

Hypothesis 1: Job demand variability moderates the effects of challenge demand level such that more variable challenge demands have a stronger positive effect on (a) strain and (b) engagement than more stable challenge demands.

Hypothesis 2: Job demand variability moderates the effects of hindrance demand level such that more variable hindrance demands have a stronger (a) positive effect on strain and (b) negative effect on engagement than more stable hindrance demands.

Method

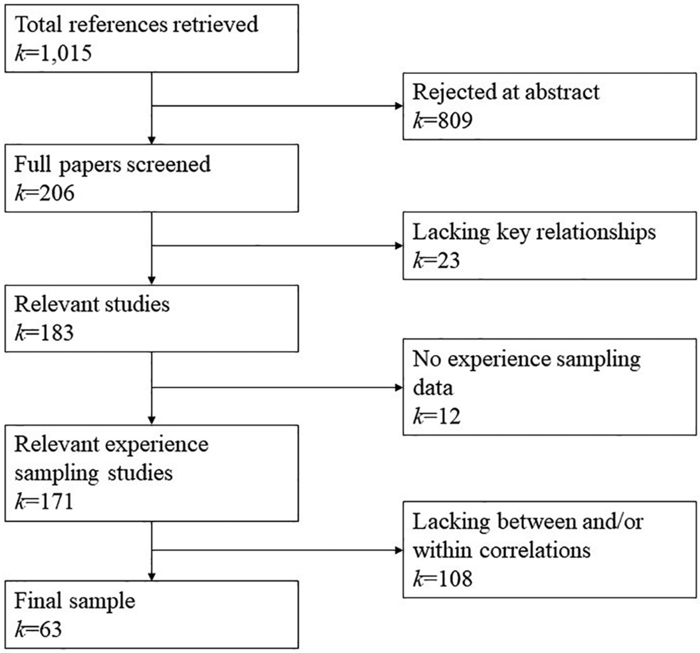
We examine our hypotheses by meta-analyzing studies that have employed ESM involving frequent (e.g., daily) measurements of demands, strain, and engagement. These ESM studies are ideal for answering our research questions because they separate and report the total variance of challenge and hindrance demands into the portion that lies between persons (i.e., the variance in person-level mean demands) and the portion that lies within persons (i.e., the variance represented by each person's deviation from his or her typical levels). The between-person estimate, as an average of multiple repeated measures over time, captures the typical level of demands associated with each individual, and the within-person estimate reflects each observation's variable demands in terms of the average within-person deviation from each person's typical (i.e., mean) level of demands. The amount of job demand variability can be easily computed and reported as an ICC or simply as the percent variance attributed to each level. After reviewing and coding studies using the procedures described below, we conducted a meta-analytic path analysis to test our hypotheses. A meta-analytic approach to our research question offers enhanced external validity by examining job demand variability across job types, organizational contexts, and industries. This is central to our question about whether there is a difference in the between-person effects of challenge and hindrance demands in contexts where the levels of job demands are more stable versus more variable. A meta-analytic approach also allows us to update the existing between-person job demands meta-analyses by including only ESM studies, thus offering greater temporal sensitivity and ecological validity (Beal, 2015).

Search and Coding Procedures

We searched the PsycINFO and Web of Science databases through February 2018 for relevant articles. After reviewing the search terms used in previously published meta-analyses focused on job demands, job resources, strain, and engagement (Alarcon, 2011; Crawford et al., 2010; Nahrgang et al., 2011; Podsakoff et al., 2007), we combined four sets of search terms using Boolean operators such that an article would need to contain at least one term from each set in order to be returned. The first set of search terms (*demands, resources, workload, role ambiguity, role conflict, control, autonomy, stressor*) encompassed terminology regarding job demands and resources. The second set of terms (*strain, engagement*) contained our outcome terms. A third set (*experience sampling, event sampling, momentary assessment, diary study, everyday experience method, daily diary, within-person, between-person*) included terms in order to identify studies that explored both within- and between-person relationships. Finally, the fourth set of search terms (*job, work, employment, organization*) was added in order to restrict the sample to studies focused on an organizational context. This search resulted in 182 references from Web of Science and 123 references from PsycINFO. To ensure that our search was as expansive as possible, we also used the same search terms in a Google Scholar search performed with an authenticated ProQuest subscription in May 2019. This additional search produced 710 references, for a total of 1,015 potential studies.

In order to be included in our final sample, studies needed to have used ESM to assess job demands, resources, strain, or engagement within individuals. Further, studies needed to have reported within- and between-individual correlations. Consistent with Gabriel et al. (2019), we found rather incomplete reporting among prior studies—108 of 171 studies at this stage did not report both within- and between-person correlations. After filtering the initial search results based on our inclusion criteria, our final dataset coded samples from 63 manuscripts for between- and within-person correlations as well as sample size and variance attributable to the within-person level. Figure 1 illustrates the number of studies excluded in each step.

Figure 1
Summary of Study Selection and Exclusion



The first and second authors jointly developed the decision rules and coding procedure for all articles that met inclusion criteria. The first author then independently coded the studies, discussing and resolving through consensus with the second author any questions that arose. As a check for accuracy, the second author independently coded approximately 14% of the studies (k = 9); no discrepancies were recorded. Following previous meta-analyses (Crawford et al., 2010; LePine et al., 2005), demands were categorized based upon the two-dimensional challenge-hindrance framework (Cavanaugh et al., 2000). Challenge demands thus included measures of workload, time pressure, felt challenge, skill demands, work pressure, and work demands. Hindrance demands included measures of conflict, constraints, customer conflict, hassles, incivility, and role conflict. Relying on prior between-person meta-analysis (LePine et al., 2005), measures of strains included depletion, distress, emotional exhaustion, exhaustion, fatigue, irritation, strain, and stress. We calculated composites (Schmidt & Hunter, 2015) for studies that employed more than one measure of demands, strain, or engagement (k = 13).

We also coded job resources as aspects of a job that are functional in achieving work goals (Demerouti et al., 2001). Examples of job resources are workplace social support (from coworkers, supervisors, or the organization), autonomy, job control, role clarity, job resources, receiving help, positive events, and positive collegial interactions. We controlled for job resources given that they may also influence strain and engagement (Crawford et al., 2010; Gonzalez-Mulé & Cockburn, 2017; Karasek, 1979).

Meta-Analytic Technique and Assumptions

We analyzed the data following procedures for bare-bones meta-analysis (Schmidt & Hunter, 2015). In bare-bones meta-analysis, sample-weighted effect sizes are computed for each relationship, which approximates weighting by the inverse of each study's sampling error (Schmidt & Hunter, 2015). However, these procedures do not correct for measurement error or other statistical artifacts. We chose not to correct effect sizes for these artifacts due to the ambiguity about how measurement error affects within- and between-person correlations as compared to single-level effect sizes. In classical psychometric theory, the underlying construct is theorized to be stable over time. In this approach, measurement errors are assumed to be uncorrelated because within-person change represents random error (Le, Schmidt, & Putka, 2009). However, in the ESM approach, within-person change is viewed as substantive in that constructs are expected to vary over time. This can lead to errors (both transient errors and specific item or scale errors) that are correlated over time. For this reason, reliability corrections that are common in traditional between-person meta-analyses may not be appropriate in ESM given that between-person correlations are formed from aggregated observations (see Lüdtke, Marsh, Robitzsch, Trautwein, Asparouhov, & Muthén, 2008, for further discussion of reliability of reflective constructs). Finally, there are multiple techniques researchers in primary studies could use to calculate reliability from multilevel ESM data. Recommendations in this area are relatively nascent (Geldhof, Preacher, & Zyphur, 2014), and the technique used to compute reliability in primary studies was rarely reported, introducing uncertainty into potential corrections we could have devised.[2]

After computing average sample-size weighted correlations, we conducted full information meta-analytic structural equation modeling (FIMASEM) using procedures described by Yu, Downes, Carter, and O'Boyle (2016)[3] analyzing the data using the R package 'lavaan' 0.6-3 (Rosseel, 2012). Specifically, we used the correlation matrix in Table 1 as the input matrix for path analysis, setting the variances of each construct to 1. We followed the recommended procedure of entering the harmonic mean of sample sizes across the relevant cells ($N_{between}$ = 1,451) as the sample sizes. Two advantages of FIMASEM over point-estimate meta-analytic structural equation modeling (MASEM) (Viswesvaran & Ones, 1995) are as follows: (1) FIMASEM does not rely upon null hypothesis significance testing, and (2) FIMASEM estimates the heterogeneity of effect sizes in a structural model. Therefore, in lieu of emphasizing statistical significance, we interpret our results by relying on the distribution of effect sizes.

This procedure makes two key assumptions that could influence our results and should thus be explicitly discussed. First, we assume correlations in primary studies were calculated using WABA techniques (Dansereau, Alutto, & Yammarino, 1984). This calculation is identical to the between-person correlation of the aggregated Level 2 means for X and Y. We excluded studies where the reported procedure for computing correlations was not consistent with this approach (e.g., averaging across individuals' raw score daily correlations). However, we included studies that did not fully disclose the procedure for computing the within-person correlations as long as they differentiated the between-person from the within-person correlations. Second, we assume our results are not a function of our study weighting strategy. Ideally, correlations in meta-analysis give more weight to more precise studies. In a traditional meta-analysis, this is usually accomplished using inverse variance (Hedges & Olkin, 1985) or sample-size weighting (Schmidt & Hunter, 2015). Based on research examining weighting

Table 1
Meta-Analytic Results of Within- and Between-Person Correlations

	Job Challenge Demands	Job Hindrance Demands	Job Resources	Psychological Strain	Job Engagement
Job Challenge Demands					
$\bar{r}$, SDρ		.24, .17	.08, .27	.20, .15	.11, .18
95% CI		(.14, .33)	(.08, .24)	(.13, .28)	(.01, .21)
80% CV		(.00, .47)	(–.28, .44)	(.00, .41)	(–.13, .35)
k; N		16; 8,859	13; 6,133	19; 14,807	16; 6,342
Job Hindrance Demands					
$\bar{r}$, SDρ	.38, .14		−.02, .12	.22, .21	−.10, .07
95% CI	(.29, .46)		(–.10, .07)	(.11, .33)	(–.17, –.02)
80% CV	(.19, .56)		(–.18, .15)	(–.05, .49)	(–.19, .00)
k; N	16; 1,862		10; 8,788	17; 10,917	7; 2,704
Job Resources					
$\bar{r}$, SDρ	.10, .34	−.14, .19		−.11, .15	.31, .12
95% CI	(–.11, .31)	(–.29, .01)		(–.20, –.02)	(.25, .37)
80% CV	(–.36, .56)	(–.40, .12)		(–.31, .10)	(.15, .46)
k; N	13; 1,525	10; 1,283		14; 9,309	21; 10,178
Psychological Strain					
$\bar{r}$, SDρ	.27, .25	.38, .24	−.20, .16		−.26, .18
95% CI	(.15; .40)	(.25, .51)	(–.30, –.09)		(–.38, –.15)
80% CV	(–.05, .60)	(.06, .71)	(–.41, .02)		(–.50, –.02)
k; N	19; 2,053	17; 2,037	14; 1,504		12; 6,401
Job Engagement					
$\bar{r}$, SDρ	.11, .21	−.12, .10	.36, .19	−.37, .14	
95% CI	(–.01, .23)	(–.25, .01)	(.26, .45)	(–.47, –.26)	
80% CV	(–.17, .38)	(–.27, .02)	(.11, .61)	(–.56, –.18)	
k; N	16; 1,446	7; 751	21; 1,984	12; 1,333	

Note: Correlations below the diagonal are between-person; correlations above the diagonal are within-person.

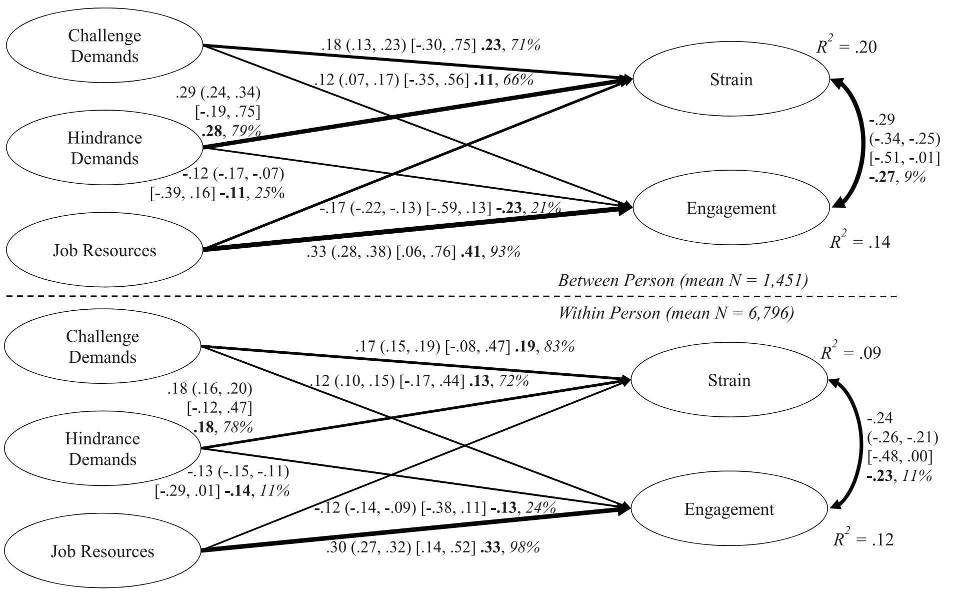
procedures (Field, 2005; Hall & Brannick, 2002) and given that our meta-analysis is based on correlations (Brannick, Yang, & Cafri, 2010), we employ sample size weighting.

Results

Bivariate Meta-Analysis

Table 1 reports the results of the meta-analysis with all data included. We conducted a sensitivity analysis (Field, Bosco, Kepes, McDaniel, & List, 2018) to ensure results were robust to outliers; removing outliers did not substantially affect results in Table 1.[4] Challenge and hindrance demands were positively related to psychological strain between individuals ($\bar{r}$ = .27 and .38, respectively), though both effects exhibited significant true score heterogeneity (SDρ = .25 and .24, respectively). Consistent with prior research, results from the meta-analysis also showed that challenge demands had a positive, but weak, association with engagement ($\bar{r}$ = .11, SDρ = .21), and hindrance demands had a negative, but weak, association with engagement ($\bar{r}$ = –.12, SDρ = .10). Again, both effects exhibited meaningful true score heterogeneity.

Figure 2
Meta-Analytic Structural Equation Model at Between- and Within-Person Levels (All Data Included)



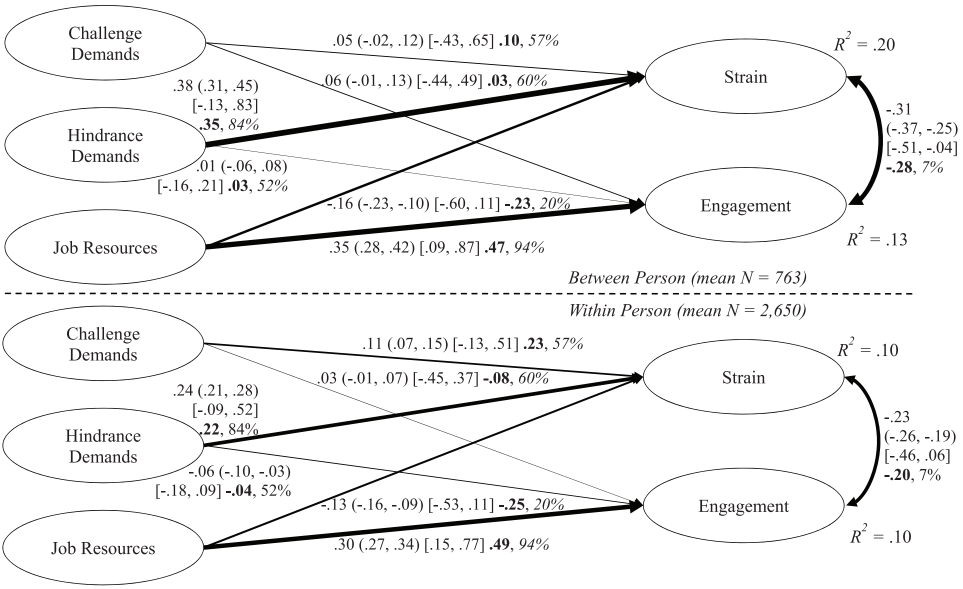
Note: All coefficients are standardized. Estimates reflect point estimate MASEM, with 95% confidence intervals in parentheses; 80% credibility intervals in brackets; FIMASEM mean beta in bold; FIMASEM percent positive betas italicized. Arrow weights are scaled to be proportional with the MASEM point estimate.

Meta-Analytic Path Analysis

Figures 2-4 present a combination of results using point estimate MASEM (Viswesvaran & Ones, 1995) and FIMASEM (Yu et al., 2016). First, the point estimate for each path, along with the 95% confidence interval for that point estimate, is reported using point estimate MASEM. FIMASEM bootstraps a population of studies representative of the ρ and SDρ matrix (i.e., Table 1). Consistent with Yu et al. (2016), we report the average β, the 80% credibility interval for β, and the percent of iterations where $\beta > 0$. In lieu of reporting statistical significance, we rely on these statistics to determine the extent to which the data are consistent with our hypotheses. To better interpret the magnitude of each effect size, lines in the figures are weighted proportionally such that larger effect sizes are depicted by wider arrows.

MASEM results based on all available data are presented in Figure 2. Findings are again consistent with those from previous analyses focused on between-person studies of job demands, strain, and engagement. In particular, challenge demands were positively related to strain (β = .18, 80% CV [–.30, .74], 71% positive) and had a weaker positive relationship with engagement (β = .12, 80% CV [–.35, .56], 66% positive). Hindrance demands were positively related to strain (β = .29, 80% CV [–.19, .75], 79% positive) and negatively related to engagement (β = –.12, 80% CV [–.39, .16], 75% negative). The results were similar to those from Crawford et al. (2010), though the effects of both types of demands on engagement were smaller than in Crawford et al. (note this is likely due in part to measurement error corrections Crawford et al. could perform).

Figure 3
Meta-Analytic Structural Equation Model at Between- and Within-Person Levels (Low Within-Person Variability Demands)



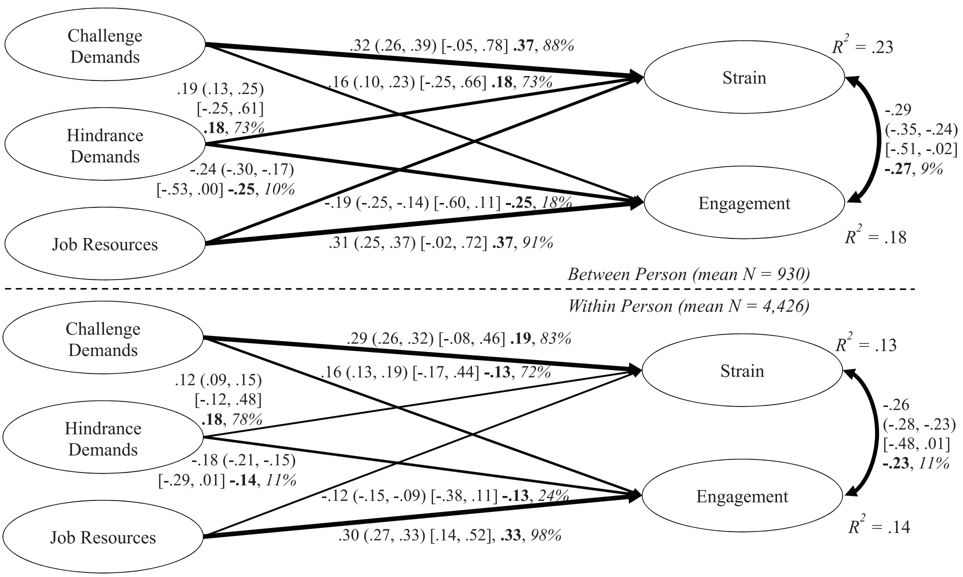
Note: All coefficients are standardized. Estimates reflect point estimate MASEM, with 95% confidence intervals in parentheses; 80% credibility intervals in brackets; FIMASEM mean beta in bold; FIMASEM percent positive betas italicized. Arrow weights are scaled to be proportional with the MASEM point estimate.

We then turned our attention toward the moderating influence of within-person variability. The correlation between the ICC (within-person) of challenge demands and the effect size was .32 for strain and .09 for engagement. For hindrance demands, the correlation between the ICC (within-person) and the effect size was .01 for strain and –.45 for engagement. We divided the effect sizes into low- and high-variability subsets splitting at the mean ICC of challenge demands (mean = .48) and hindrance demands (mean = .48). When ICCs were not reported, we filled in the average ICC for that specific type of challenge or hindrance demand (e.g., mean within-person ICC for the challenge demand of time pressure = .56). When multiple measures of demands were used, we used the average within-person variability across the multiple measures to determine the within-person job demand variability.

Note that our subgrouping strategy is at the effect-size level and therefore splits the effect sizes in Table 1 independently for each cell. This means that it is possible for a single study to have effect sizes that feed into the low variability *and* the high variability subgroups (e.g., if challenge demand variability were high and hindrance demand variability were low). This occurred for five studies; results with these five studies removed are available from the first author.

Figure 3 displays results based on studies where demands had lower levels of within-person variability. Figure 4 displays results based on studies where demands had higher levels of within-person variability. Note that in each figure we include both the between- and

Figure 4
Meta-Analytic Structural Equation Model at Between- and Within-Person Levels (High Within-Person Variability Demands)



Note: All coefficients are standardized. Estimates reflect point estimate MASEM, with 95% confidence intervals in parentheses; 80% credibility intervals in brackets; FIMASEM mean beta in bold; FIMASEM percent positive betas italicized. Arrow weights are scaled to be proportional with the MASEM point estimate.

within-person levels of analysis. Our hypotheses are restricted to the between-person level of analysis, and the within-person level of analysis is reported solely for completeness.

Hypothesis 1 proposed that the positive effects of challenge demands on (a) strain and (b) engagement would be moderated by job demand variability such that challenge demands would have stronger positive between-person effects when within-person variability is high than when within-person variability is low. As shown in Figures 3 and 4, challenge demands had a stronger positive effect on strain in studies with high within-person variability (β = .32, 80% CV [–.05, .78], 88% positive) than in studies with low within-person variability (β = .05, 80% CV [–.43, .65], 57% positive). These results were consistent with Hypothesis 1(a). Challenge demands also had a stronger positive relationship with engagement in studies with high within-person variability (β = .16, 80% CV [–.25, .66], 73% positive) than in studies with low within-person variability (β = .06, 80% CV [–.44, .49], 60% positive). Note that supplementary analyses without mean-substituting for missing ICCs (see Online Supplement) showed the difference in this effect was small. Thus, although results from the full sample are consistent with Hypothesis 1(b), this did not hold in the smaller sample without mean-substituting missing values.

Hypothesis 2(a) posited that the between-person effects of hindrance demands on strain would be more positive when the within-person variability of the demand was high rather than low. Although hindrance demands had the expected positive effect on strain, Hypothesis 2(a) was not supported given the coefficients in studies with high (β = .19, 80% CV [–.25, .61], 73% positive) and low within-person variability (β = .38, 80% CV [–.13, .83], 84% positive). Hypothesis 2(b) posited that the negative effect of hindrance demands on engagement would be stronger in studies with high within-person variability than in studies with low variability. Results showed the effect was stronger in studies with high variability (β = –.24, 80% CV [–.53, .00], 90% negative) than in studies with low within-person variability (β = .01, 80% CV [–.16, .21], 48% negative). This result was consistent with Hypothesis 2(b).

Supplementary Power Analysis

The meta-analytic technique splits primary studies into those exhibiting low versus high degrees of within-person job demand variability. One advantage of this approach over a single primary study is that it accesses available evidence across primary studies on the topic. This means the data come from a broader set of industries, organizations, and occupations, increasing the likelihood that our findings generalize to the population of workers. However, this technique to some degree separates our theory about job demand variability from our empirical test of its moderating effects. That is, our theory suggests that *individuals* with higher levels of within-person job demand variability will see stronger effects of those demands on their engagement and strain. Yet our empirical test examines whether *primary studies* with higher levels of within-person job demand variability show stronger effects than studies with lower levels of job demand variability. This represents an ecological fallacy in theorizing at the individual level but testing the data at the study level. Given the novelty of this approach to combining ESM studies, additional exploration into its inferential validity is warranted.

Overview of Procedure

We conducted a simulation to explore the validity of the method in terms of its ability to report hypothesized effects in a population where the effect is known to exist (and to detect noneffects in a population where the effect is known not to exist). To simulate these instances, we established four populations of 10,000 individuals where the moderation effect size of within-person job demand variability was zero, small, medium, and large at the individual level (i.e., Level 2). To create these population datasets, we sampled Level 2 scores (i.e., the individual-level means) from a normal distribution with a mean of 0 and standard deviation of 1. We then sampled 10 observations for each individual (consistent with a typical ESM study) from a normal distribution[5] with a mean of 0 and a standard deviation randomly selected (per individual) to be between .50 and 1.50. This procedure resulted in study-level ICC(1) values that could theoretically range from approximately 31% to 80%, with an average of about 50%. This is consistent with study-level ICC(1) values found in ESM literature (see Tables 2 and 3 of McCormick et al., 2020). The R syntax used to generate the population datasets (as well as reproduce the entire simulation) is available in an online supplement. For parsimony, the dataset examined one Level 2 independent variable (i.e., challenge/hindrance demands), one Level 2 dependent variable (i.e., strain/engagement), and one Level 2 moderator (i.e., job demand variability). Consistent with our focus on between-person effects, all variables are approached from the individual level (i.e., the within-person scores are generated merely for the purpose

Table 2
Results of Supplementary Power Analysis

	Moderation Effect Size							
	None (R^2 = .00)[a]		Small (R^2 = .01)[a]		Medium (R^2 = .09)[a]		Large (R^2 = .16)[a]	
Meta-ICC[b]	ME Approach	MA Approach	ME Approach	MA Approach	ME Approach	MA Approach	ME Approach	MA Approach
Estimated from studies (.18[c])	.026	.008	.045	.488	.118	.974	.192	.990
Large (.10)	.025	.006	.079	.488	.362	.965	.603	.987
Medium (.07)	.026	.016	.135	.343	.719	.939	.924	.978
Small (.03)	.024	.024	.228	.117	.920	.628	.990	.738
None (.00)	.025	.014	.400	.009	.994	.023	.999	.037

Notes: ME = mixed effects technique; MA = meta-analysis technique. Data are based on 1,000 bootstrapped meta-analyses (14 primary studies with 100 participants per study [10 observations per participant] included in each meta-analysis). We used 95% confidence intervals of meta-analytic estimates to evaluate statistical significance.
[a]Effect sizes are based on Pseudo R^2 change of adding the interaction term to the model.
[b]Percent figures in parentheses indicate the percent of variance in within-person variability (i.e., ICC1 of study) in the meta-analysis.
[c]Meta-ICC cannot be computed without primary data for all studies; it is estimated based on matching the observed variance in study ICCs in the field to the equivalent variance in the simulation.

of calculating within-person variability). Once the population dataset was created, we bootstrapped 14 primary studies of 100 individuals from the population for each meta-analysis. We chose 14 primary studies as that approximated the *k* in our field meta-analysis.

A key aspect of sampling primary studies involves the extent to which the primary studies randomly sample across the full population of within-person job demand variability. In our meta-analysis, Level 1 observations are nested within individuals (Level 2), who are then nested within primary studies. If primary studies all randomly draw from the same full population, then we would expect them to on average have the same amount of within-person variability. If there are nonrandom factors that influence the sampling from the population, then we would expect studies to differ in the average amount of within-person variability they observe (because primary studies would capture different, narrower slices of the distribution of within-person variability). This can be thought of as a meta-ICC (as it represents an ICC of ICCs) or the percent of variability (in within-person variability) that is attributable to the study level.

At the primary study level, there are several reasons to believe that sampling is not random across within-person job demand variability. For example, some samples access populations within the same job, within the same organization, or from the recruiting source; it is plausible that any single primary study is restricted in the amount of within-person job demand variability it can capture. There is also recent evidence suggesting that study design can influence the amount of study-level within-person variability observed in a given study (Podsakoff, Spoelma, Chawla, & Gabriel, 2019). To more fully understand how study-level variance in within-person variability affected the validity of our analytic technique, we explicitly modeled different conditions of study-level variance in within-person variability.

To accomplish this, the simulation randomly chose a within-person standard deviation, then limited further sampling for that primary study to individuals within a condition-defined range of within-person variability. The narrower the sampled range, the more within-person variability varied between primary studies, and the larger the meta-ICC. The wider the range, the more within-person variability was similar across studies (because each primary study sampled from a wider range of the distribution), and the smaller the meta-ICC. We manipulated primary-study sampling in this way to achieve conditions where the meta-ICC was .00, .03, .07, and .10.

Note that the raw data from all primary studies are required to compute the meta-ICC. Because we did not have access to raw data from all included studies, the meta-ICC for our field meta-analysis could not be computed. However, we could estimate the meta-ICC by comparing the variance in study-reported ICCs to the variance of ICCs in the simulated primary studies. This comparison resulted in a meta-ICC of .18, which was larger than the conditions we a priori developed. For comparison, we included this fifth level of meta-ICC in the simulation.

In sum, the two factors (effect size and meta-ICC) with four and five levels, respectively, produced 20 simulation conditions. Each of four population distributions (corresponding to the various effect sizes of moderation) were sampled five separate times (corresponding to each meta-ICC value). For each condition, we conducted 1,000 meta-analyses with a *k* of 14 primary studies (using the "psychmeta" [Dahlke & Wiernik, 2019] and "multilevel" packages [Bliese, 2002]). Once the 14 studies were sampled, we conducted both a mixed-effects analysis (that would be used in a primary study) as well as a meta-analysis (mirroring our above approach).

For the mixed-effects technique, we freed the random intercept for each individual, person-level centered the Level 1 predictor, and modeled the Level 2 interaction between the aggregated mean and the standard deviation of Level 1 observations. We then examined the statistical significance of the interaction term in the primary study. For the meta-analytic approach, we first computed a mean ICC across the 14 studies, then dichotomized studies into high and low variability based on that mean value. Note that dichotomizing studies into either high or low variability results in a loss of information and that a continuous-ICC approach would be preferred. However, such techniques (i.e., continuous moderator analysis in MASEM) have not been developed, and our field study was limited to dichotomization. To keep the simulation and field study parallel, we used subgroup moderator analysis splitting the 14 studies into high- and low-variability groups before computing confidence intervals. Because power analyses require a statistical significance test, we computed the percent of the 1,000 meta-analyses where the 95% confidence intervals for the two subsamples did not overlap.

Supplementary Analysis Results

Results presented in Table 2 report the percent of null hypotheses that were rejected using either the meta-analytic or the mixed-effects approach. Table 2 illustrates several key points for understanding the validity of our proposed MASEM technique. For example, column 1 of Table 2 shows that the Type I error rates of the meta-analytic technique range from .006 to .024; this is similar to the Type I error rates of the mixed-effects model approach (ranging from .024 to .026). If the results we detected were merely a function of the methodological technique of splitting studies into high and low variability, we would expect inflated Type I errors (i.e., the technique would report an effect when none exists). The low Type I error rates thus offer evidence that the moderation effect is not exclusively due to the steps involved in the meta-analytic approach (e.g., the way in which Level 2 scores are aggregated, the dichotomization into high- and low-variability studies, the examination of hypotheses at the study, rather than the individual, level).

Another key point from the simulation focuses on whether our proposed meta-analytic approach detects known effects. To explore this question, we compare the statistical power of the mixed-effects approach that would be utilized in a primary study against our proposed meta-analytic approach. The remaining columns of Table 2 show that the answer to this question depends on the meta-ICC (i.e., the degree to which primary studies vary in the amount of within-person variability they observe). For example, if all the studies are randomly drawn from the same population (and therefore observe the same amount of within-person variability), the meta-ICC is small and the mixed-effects approach offers more statistical power. The bottom row of Table 2 illustrates that when the meta-ICC is 0 (i.e., there is no study-level variance in ICCs), the meta-analytic approach is unable to detect a small (Pseudo R^2 = .01), medium (Pseudo R^2 = .09), or even large (Pseudo R^2 = .16) effect (power = .009, .023, and .037, respectively). In contrast, the mixed-effects approach is quite good across these conditions, exhibiting nearly perfect power when the effect size was medium (power = .994) or large (power = .999).

However, as the meta-ICC increases, the meta-analytic approach offers more statistical power. In conditions where the meta-ICC was .03 (i.e., 3% of the variance in within-person variability was between studies), the mixed-effects approach had sufficient power to detect a medium or large effect (power = .920 and .990, respectively). In these same conditions, our proposed meta-analytic technique lacked sufficient power (.628 and .738, respectively). However, as the meta-ICC increased to .07, the power of our proposed meta-analytic technique exceeded the power of the mixed-effects technique. This was the case for small (Pseudo R^2 = .01), medium (Pseudo R^2 = .09), or large (Pseudo R^2 = .16) moderation effects (meta-analytic power = .343, .939, and .978, respectively; mixed-effects power = .135, .719, and .924, respectively). Interestingly, as the meta-ICC increased, the power of the meta-analytic approach continued to increase, whereas the power of the mixed-effects approach decreased. In conditions where the meta-ICC was .18 (as we estimated it to be in our above meta-analysis), the technique showed power of .488, .974, and .990 to detect a small, medium, and large effect (respectively). In sum, the supplementary analysis indicates that the appropriateness of each technique depends on the degree of heterogeneity in study samples, with the meta-analytic technique offering more power in cases of greater heterogeneity such as what we observe in the meta-analysis.

Discussion

Our focus in this study was to inform theory and practice by highlighting job demand variability as a contingency factor that determines how employees are affected by job demands. Extending existing research, which has adopted either a between- or within-person perspective in isolation, we integrated both perspectives to posit that between-person differences in within-person job demand variability would moderate the effects of job demands on employee strain and engagement. Integrating arousal theory with extant theory on job demands, we hypothesized that more stable demands should consume fewer personal resources and have generally weaker effects on strain and engagement than more variable job demands.

In line with our expectation, challenge demands had stronger positive effects on strain when employees encountered higher levels of within-person challenge demand variability versus when they encountered lower levels of challenge demand variability. These findings are consistent with our expectation that when higher levels of challenge demands are more variable, they present a more significant tax on personal resources (and therefore produce more strain) than do job demands that are more stable across employees' work experiences. Similarly, we found that the relationship between hindrance demands and engagement was stronger under conditions of more variable (vs. more stable) job hindrance demands. This suggests that hindrance demands reduce engagement only when they are also variable over the course of employees' work experiences. When high levels of hindrance demands are more stable over employees' experiences, they have a nonsignificant relationship with engagement.

In terms of the relationship between challenge demands and engagement, the results were mixed. In the full sample of studies, results were consistent with our hypothesis that challenge demands would have a stronger relationship with engagement under conditions of high challenge demand variability rather than low variability. However, in a robustness check examining a subsample of studies (i.e., those without missing ICC values), the results were not consistent with this hypothesis. Thus, although the evidence in the full sample was consistent with our hypothesis, we caution that study design decisions may have influenced this conclusion. Future research with more consistently complete data to assess this proposition would be welcome.

Additionally, the effects of hindrance demands on strain were not as expected, as hindrance demands had a significant positive effect on strain regardless of the extent to which those hindrance demands varied over the course of employees' experiences. Several post hoc explanations for this finding are plausible. As one possibility, individuals' use of different coping strategies could explain why hindrance demands' effects persist even when they are relatively stable. Researchers have found challenge demands to initiate "an active or problem-solving style of coping (e.g., strategizing, increases in effort)" (Crawford et al., 2010: 837) and have posited that this "active problem-solving coping style is more adaptive" (Wallace, Edwards, Arnold, Frazier, & Finch, 2009: 256). It may be that stable challenge demands do not induce strain because employees have taken an active role in overcoming or "getting used to" them.

The extent to which hindrance demands also incite such active responses is less clear. Some researchers have found that hindrance demands lead to active responses such as job crafting and thus reduce strain (Bakker, Demerouti, & Sanz-Vergel, 2014; Tims, Bakker, & Derks, 2012, 2013). This may be particularly true for certain individuals (e.g., high in proactive personality; see Tims & Bakker, 2010). Yet other researchers argue the hindrance demands tend to "trigger negative emotions . . . and induce a passive or emotional style of coping (e.g., rationalization, withdrawing from the situation)" (Crawford et al., 2010: 837). Pearsall, Ellis, and Stein (2009: 21) described that employees can "view hindrance stressors as unchangeable" and "abandon attempts to improve the situation." Thus, if this passive coping strategy is the one chosen in response to hindrance stressors, it could prevent employees from overcoming, adapting, or otherwise changing their stable hindrance demands in such a way that reduces strain.

An alternative (but still compatible) interpretation of our findings questions the ability of arousal theory to fully explain job demand variability; it may be that our broad application of arousal theory provides insufficient nuance into how hindrance demands affect individuals. Through an arousal theory lens, any deviation from "typical" should command more attention and thus exert a stronger effect on individuals' strain and engagement. Yet it may be that not all forms of job demand variability are equal. Because employees learn, cope, adapt, and gain skills through experience, it is plausible that the predictability (i.e., the degree to which employees can anticipate demands in advance), routineness (i.e., how common it is for employees to experience deviations from a characteristic level of demands), and novelty (i.e., the degree to which demands have been experienced before) of the job demand variability may influence the role it plays in employees' job demand experience. This may manifest as different sources of variability leading to different effects (for example, if hindrance demand variability is more routine, on average, it may have weaker moderating effects). In line with this point, job demand fluctuations that are small but experienced quite frequently by employees may have different implications than large job demand fluctuations that are experienced infrequently. As we discuss in more detail below, we were limited in our ability to test these factors due to the number of available studies examining different sources of job demand variability, time intervals, and recovery.

More broadly, however, it may be that arousal theory offers a limited explanation of the effects of job demand variability or that other more specific contingencies need to be accounted for when examining job demands variability effects. Although our research demonstrates that job demand variability influences individuals' job demand experience, the available evidence at present suggests this does not occur with both types of job demands and that future research extending our understanding of the effects of job demand variability is warranted.

Theoretical and Practical Contributions

The primary contribution of our study is to introduce the notion of job demand variability and to assess the way in which it might influence between-person relationships between job demands and the relevant motivational outcomes of strain and engagement. Job characteristics theory (Hackman & Oldham, 1975) was implicitly sensitive to the idea that variability (e.g., skill variety) would lead people to higher levels of felt meaningfulness and responsibility. Yet research building on the job characteristics model has primarily fallen in one of two distinct silos: between person or within person. Our theory, rooted in Berlyne's (1960) arousal theory, considers that the between-person effects of job demands depend upon how characteristic those job demands are of that worker's experience. In short, we highlight that people differ in how much their day-to-day experiences vary, leading to between-person variation in outcomes. This integrates the previously separated within- and between-person perspectives on job demands. Further, asking this initial question opens the door to multiple avenues of theory with respect to job demand variability. As we note above, different forms of job demand variability (varying in their predictability, novelty, and routineness) may have different motivational effects. Future research into *how* job demands vary—and contrast against workers' typical demands—would be valuable to build upon our understanding of how job demands affect workers.

Not only does this new boundary condition advance the job demands literature, but it also opens the door to a new way of thinking about variability in other areas of management research. Although within-person research across organizational behavior has recognized that phenomena vary over time (McCormick et al., 2020), the within-person literature focuses on immediate (and dynamic) outcomes. Our research employs a new perspective in that it integrates the within- and between-person perspectives by asking whether workers who have more variable experiences have different between-person outcomes than workers who have more stable experiences. This framework offers a template that can be used to advance other areas of management, such as leadership (e.g., leader behaviors that are stable and consistent may have different effects than behaviors that fluctuate), personality (e.g., employees who are stable and consistent in their expressions of conscientiousness may have different performance than those who are variable in those behaviors), or person-organization fit (e.g., outcomes of employees' perceived fit may differ depending on whether that fit is experienced consistently and regularly vs. only on certain variable occasions). More broadly, our research introduces a new kind of research question and provides empirical guidance on how this kind of research question can be tested.

Our study also contributes to the literature on job demands through our replication of Crawford et al.'s (2010) findings. Although the primary studies are different from those in Crawford et al., the effects were very similar. For this particular model, ESM and traditional between-person study designs have yielded similar results. Having this information can be valuable for future researchers and practitioners who seek to determine whether the added expense and effort of an ESM study design is warranted (see McCormick et al., 2020).

Regardless of whether our findings are disseminated to entrepreneurs who are trying to optimize their outcomes, employees crafting their own jobs, or managers and human resource professionals assessing the features of jobs, a key takeaway for practitioners is to recognize that the effects of job demands on strain and engagement tend to be different when those demands are more variable as compared to when they are more stable. Thus, to get the most out of themselves or others at work, workers need to look beyond the types of demands faced to account for the variability of demands when considering people's cognitive and affective responses to their work. Although specific examples of how a manager would deal with this information depend on the category of demand (challenge vs. hindrance) and outcome of interest (strain vs. engagement), one such example follows. If a manager is alarmed at having a set of disengaged employees and recognizes that those employees face a lot of variability in the situational constraints they encounter (a hindrance demand), then the manager could seek to alter resource allocations or expectations in order for employees to become accustomed to a more consistent level of constraints. Such consistency would be expected to mitigate the harmful effects of this hindrance demand on employee engagement.

Our study also offers a methodological contribution with a potential for broader impact beyond the job demands domain of research. Research in many areas can benefit from this method by examining the extent to which the variability or stability of phenomena facilitates or limits its effects. This expands on our understanding of dynamic phenomena to consider not only the immediate temporal effects of individuals' states, situations, and behaviors but also the way in which individuals' varying circumstances influence their typical experiences.

By articulating the meta-ICC and simulating its importance in our hypothesis, we offer a springboard for future research using a similar model. Researchers looking to understand individual phenomena in more stable versus more variable contexts are encouraged to employ meta-analysis when the meta-ICC (and thus the variance in study-reported ICC) is high.

Limitations and Future Research

Although our meta-analytic approach offers a number of strengths, it also comes with limitations. For example, we were not able to test for additional mediators and moderators of the focal relationships in our model, and we were not able to conduct a robust test of causal ordering. Although our meta-analytic approach makes a contribution in how it uses variability reported in primary studies, this requires further study. For example, we imputed the relevant mean ICC (e.g., workload, time pressure, conflict) for studies that did not originally report ICC values ($k = 11$). We conducted a supplementary analysis removing all studies not reporting the necessary ICC value(s) to understand how this influenced our results (available in an Online Supplement). Conclusions in this supplementary analysis were similar for Hypotheses 1(a) and 2. As noted above, conclusions regarding Hypothesis 1(b) differed. Additionally, we coded studies that reported within- and between-person correlations that we believed were computed by person-centering the Level 1 scores. However, there are multiple techniques scholars use to calculate these correlations, and not all studies fully disclosed the procedure they used. Thus, it is plausible that some within-person correlations were computed in different ways. Because our research is focused on Level 2 relationships, this limitation is mitigated but is an area for future research.

One additional area of future research involves the extent to which job demand variability should be conceptualized as a distinct construct from job demand level. In many jobs, the reason a job demand is experienced as demanding is because the demand is variable in some way. In this sense, it may be the case that workers experience variable job demands as automatically being high in demand level (and, in contrast, stable job demands as being rather low in demand level). From this perspective, the distinction between the level and variability of job demands is less clear: The constructs could be so highly related that the distinction would not be important in describing how workers experience job demands. We were limited in our ability to test this proposition, but note that the correlations between reported ICC and the mean level of job demands were .44 for challenge demands and –.05 for hindrance demands. These correlations should be interpreted with care given they are at the study level instead of the individual level, yet they appear to indicate a distinction between variability and overall level of job demands.

Along these lines, future research should employ variability as a part of the cognitive appraisal processes involved as individuals experience variable job demands (e.g., LePine, Zhang, Crawford, & Rich, 2016; Webster, Beehr, & Love, 2011). In line with our recommendations regarding novelty, routineness, and predictability, it may be that workers interpret some forms of variability as creating higher demand levels. For example, it may be that job demand variability induced through novelty is experienced as a high level of job demands because it is new (and therefore variable). However, job demand variability induced through regular routines may feel like a low job demand even though variable. Viewed as a cognitive appraisal process, it may be that certain kinds of job demand variability are better conceptualized as distinct from the demands themselves, whereas other kinds are better conceptualized as part-and-parcel of what it means for a job to feel demanding in the first place.

There are several additional areas for future research. Arousal theory provides a sound basis for expecting job demand variability to moderate the effects of job demand level regardless of individuals' prior experiences and expectations. Yet data in our sample—typically comprised from measures of incumbent participants over a 2-week period—did not specify employees' expectations of job demand variability or the degree to which job demands were completely novel to employees. Employees certainly had varied experiences and expectations leading into those 2-week periods that may have influenced how they responded to job demand variability. These prior experiences, the expectations they create in the minds of individuals, and the way those expectations manifest in subsequent work episodes merit further inquiry.

Another noteworthy future direction would be to better specify how the timing of job demand variability affects employee's reactions to those job demands. The bulk of studies in our sample examined job demands at a daily level, conforming to a normative view of a single 24-hour period as an episode of work. But this overlooks the fact that job demands vary substantially over the course of a single day. Different workdays may exhibit very similar levels of demands and have little day-to-day variability—but meaningful within-day variability. This may offer one explanation for the fairly wide credibility intervals we observed: In different jobs, organizations, or industries, the salient performance episode may take a different time frame than is captured in primary studies. Our approach potentially introduces some noise into our analysis that could have affected our conclusions. Future research is needed to understand how job demand variability across different timeframes affects worker outcomes. A starting point for shorter time horizons might be the literature on task switching (e.g., Kiesel et al., 2010), which describes how individuals cognitively approach rapidly changing environments.

More broadly, we were limited in our ability to explore how the nature of job demand variability may have played a role in its effects. As we note above, this includes the timing of the variability but also its predictability (i.e., individuals might respond differently to unexpected rather than expected variability), routineness (i.e., the frequency of variability), and novelty (i.e., the extent to which demands are new to incumbents). In a related vein, we were also limited in our ability to capture the shape, magnitude, and direction of variability. We explicitly conceptualized job demand variability as the overall amount of variation, but variability can take many forms; a similar level of variability can be produced by a pattern of small but repeated deviations or by a pattern having one major deviation (see Harrison & Klein, 2007, for a parallel discussion). Our meta-analysis is limited in its ability to differentiate between these variability patterns. However, future research would be welcome, perhaps adopting the lens of peak/end rules to job demand variability (e.g., Kahneman, Fredrickson, Schreiber, & Redelmeier, 1993) to determine whether certain days have an outsized influence on individuals' motivation and attitudes. More nuanced theorizing of how job demand variability affects individuals over time (e.g., comparing variability that occurs within a work shift to variability that occurs from one shift to another) may produce additional insights that could inform managers' efforts.

Finally, we relied on an arousal theory (Berlyne, 1960) perspective to explain why variable job demands should influence strain and engagement more than stable job demands.

However, there are at least two alternative plausible explanations for our results. First, individuals likely experience some coping, adaptation, or skill-building as they encounter new demands. For many individuals, job demands that were once challenging eventually become routine and require fewer personal resources to accomplish. In this sense, individuals with a job with a high level of demands (without any variability in those demands) may eventually come to see that job as relatively low in demands. Our focus in the present study is on changes to job demands, but we are limited in our ability to acknowledge the changes that occur in individuals as well. Although this explanation is consistent with our hypotheses (i.e., as individuals' skills grow they experience less arousal from stable job demands), future research would be valuable.

Also, variable job demands—particularly those tied to workload—might reduce individuals' opportunity for recovery experiences (e.g., working through breaks or staying late and missing recovery time at home). Given that recovery experiences are theorized to promote personal resources (Steed, Swider, Keem, & Liu, 2019), recovery experiences should alleviate strain and facilitate engagement. Many of the studies in our sample (42 of 63) discussed recovery, but only 5 studies reported effect sizes to include recovery in our model. This limited number of studies was not sufficient for our moderation hypothesis; however, future research considering the impact of recovery on dynamic job demands would be valuable.

Conclusion

In conclusion, our work builds upon prior job demands research by demonstrating that the effects of challenge and hindrance demands on employee engagement and strain are contingent upon the variability of the demands being faced. Our perspective accounts for daily fluctuation in job demands and offers a new way to view employee demands in the workplace. Considering job demand variability as a contingency factor in frameworks of job demands offers many new theoretical and practical avenues to understand how job demands relate to employees' motivation and well-being. We encourage scholars to continue to explore the temporal complexities underlying how and when job demands affect employee outcomes and experiences.

ORCID iDs

Patrick E. Downes https://orcid.org/0000-0002-8944-020X
Brian W. McCormick https://orcid.org/0000-0002-1308-0551

Notes

1. Several models of job demands also recognize job *resources* as characteristics of a job that are functional in enhancing motivation and reducing strain (e.g., Bakker & Demerouti, 2007). Our focus is on the challenge-hindrance framework, which primarily emphasizes different types of job demands and their effects. We do include job resources in coding and analyses.

2. Results with corrections for measurement error are available by request of the first author.

3. Results using two-stage meta-analytic structural equation modeling (i.e., TSSEM; Cheung & Chan, 2005) are available in an online supplement.

4. The Field et al. (2018) sensitivity analysis is based on the R package 'metafor,' which relies on inverse variance weights and z-transformations from the Hedges and Olkin (1985) method. Relying on the Schmidt and

Hunter (2015) method, our analysis uses sample size weighting and does not transform correlations. However, because we conducted bare bones meta-analysis (i.e., no unreliability corrections) and did not observe extreme values (i.e., for our analysis, most correlations were less than .40), the differences in weighting strategies are likely to be small. Thus, we expect Field et al.'s techniques should provide reasonable estimates in terms of the sensitivity to outliers.

5. Drawing the centered within-person scores from a separate normal distribution ensured that the between-person means and centered within-person scores were orthogonal (we confirmed a zero correlation between the between- and centered within-person scores in all conditions). This prevents potential concerns of ceiling or floor effects with extreme values when means and standard deviations are correlated (Baird, Le, & Lucas, 2006).

References

Alarcon, G. M. 2011. A meta-analysis of burnout with job demands, resources, and attitudes. *Journal of Vocational Behavior*, 79: 549-562.

Baird, B. M., Le, K., & Lucas, R. E. 2006. On the nature of intraindividual personality variability: Reliability, validity, and associations with well-being. *Journal of Personality and Social Psychology*, 90: 512-527.

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., & Sanz-Vergel, A. I. 2014. Burnout and work engagement: The JD-R approach. *Annual Review of Organizational Psychology and Organizational Behavior*, 1: 389-411.

Bakker, A. B., & Sanz-Vergel, A. I. 2013. Weekly work engagement and flourishing: The role of hindrance and challenge job demands. *Journal of Vocational Behavior*, 83: 397-409.

Bargh, J. A., & Chartrand, T. L. 1999. The unbearable automaticity of being. *American Psychologist*, 54: 462-479.

Barrick, M. R., Stewart, G. L., Neubert, M. J., & Mount, M. K. 1998. Relating member ability and personality to work-team processes and team effectiveness. *Journal of Applied Psychology*, 83: 377-391.

Beal, D. J. 2015. ESM 2.0: State of the art and future potential of experience sampling methods in organizational research. *Annual Review of Organizational Psychology and Organizational Behavior*, 2: 383-407.

Beehr, T. A., Jex, S. M., Stacy, B. A., & Muray, M. A. 2000. Work stressors and coworker support as predictors of individual strain and job performance. *Journal of Organizational Behavior*, 21: 391-405.

Berlyne, D. E. 1960. *Conflict, arousal, and curiosity*. New York, NY: McGraw-Hill.

Berlyne, D. E., Borsa, D. M., Craw, M. A., Gelman, R. S., & Mandell, E. E. 1965. Effects of stimulus complexity and induced arousal on paired-associate learning. *Journal of Verbal Learning and Verbal Behavior*, 4: 291-299.

Berlyne, D. E., Borsa, D. M., Hamacher, J. H., & Koenig, I. D. 1966. Paired-associate learning and the timing of arousal. *Journal of Experimental Psychology*, 72: 1-6.

Berlyne, D. E., & Carey, S. T. 1968. Incidental learning and the timing of arousal. *Psychonomic Science*, 13: 103-104.

Bidwell, M. J. 2013. What happened to long-term employment? The role of worker power and environmental turbulence in explaining declines in worker tenure. *Organization Science*, 24: 1061-1082.

Bliese, P. D. 2002. *Multilevel modeling in R: A brief introduction to R, the multilevel package, and the NLME package*. Washington, DC: Walter Reed Army Institute of Research.

Bloom, M. 1999. The performance effects of pay dispersion on individuals and organizations. *Academy of Management Journal*, 42: 25-40.

Boswell, W. R., Olson-Buchanan, J. B., & LePine, M. A. 2004. Relations between stress and work outcomes: The role of felt challenge, job control, and psychological strain. *Journal of Vocational Behavior*, 64: 165-181.

Brannick, M. T., Yang, L.- Q., & Cafri, G. 2010. Comparison of weights for meta-analysis of *r* and *d* under realistic condition. *Organizational Research Methods*, 14: 587-607.

Cappelli, P., & Keller, J. R. 2013. Classifying work in the new economy. *Academy of Management Review*, 38: 575-596.

Cavanaugh, M. A., Boswell, W. R., Roehling, M. V., & Boudreau, J. W. 2000. An empirical examination of self-reported work stress among U.S. managers. *Journal of Applied Psychology*, 85: 65-74.

Cheung, M. W. L., & Chan, W. 2005. Meta-analytic structural equation modeling: A two-stage approach. *Psychological Methods*, 10: 40-64.

Cox-Fuenzalida, L. E. 2007. Effect of workload history on task performance. *Human Factors*, 49: 277-291.

Crawford, E. R., LePine, J. A., & Rich, B. L. 2010. Linking job demands and resources to employee engagement and burnout: A theoretical extension and meta-analytic test. *Journal of Applied Psychology*, 95: 834-848.

Dahlke, J. A., & Wiernik, B. M. 2019. psychmeta: An R package for psychometric meta-analysis. *Applied Psychological Measurement*, 43: 415-416.

Dalal, R. S., Bhave, D. P., & Fiset, J. 2014. Within-person variability in job performance: A theoretical review and research agenda. *Journal of Management*, 40: 1396-1436.

Dansereau, F., Alutto, J. A., & Yammarino, F. J. 1984. *Theory testing in organizational behavior: The varient approach*. Englewood Cliffs, NJ: Prentice Hall.

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.

Field, A. P. 2005. Is the meta-analysis of correlation coefficients accurate when population correlations vary? *Psychological Methods*, 10: 444-467.

Field, J., Bosco, F. A., Kepes, S., McDaniel, M. A., & List, S. 2018. Introducing a comprehensive sensitivity analysis tool for meta-analytic reviews. *Academy of Management Proceedings*. Advance online publication. doi:10.5465/AMBPP.2018.18291abstract

Gabriel, A. S., Podsakoff, N. P., Beal, D. J., Scott, B. A., Sonnentag, S., Trougakos, J. P., & Butts, M. M. 2019. Experience sampling methods: A discussion of critical trends and considerations for scholarly advancement. *Organizational Research Methods*, 22: 969-1006.

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

Gonzalez-Mulé, E., & Cockburn, B. 2017. Worked to death: The relationships of job demands and job control with mortality. *Personnel Psychology*, 70: 73-112.

Hackman, J. R., & Oldham, G. R. 1975. Development of the job diagnostic survey. *Journal of Applied Psychology*, 60: 159-170.

Hall, S. M., & Brannick, M. T. 2002. Comparison of two random-effects methods of meta-analysis. *Journal of Applied Psychology*, 87: 377-389.

Harrison, D. A., & Klein, K. J. 2007. What's the difference? Diversity constructs as separation, variety, or disparity in organizations. *Academy of Management Review*, 32: 1199-1228.

Hauck, E. L., Snyder, L. A., & Cox-Fuenzalida, L. E. 2008. Workload variability and social support: Effects on stress and performance. *Current Psychology*, 27: 112-125.

Hedges, L. V., & Olkin, I. 1985. *Statistical methods for meta-analysis*. New York: Academic Press.

Helton, W. S., Shaw, T., Warm, J. S., Matthews, G., & Hancock, P. 2008. Effects of warned and unwarned demand transitions on vigilance performance and stress. *Anxiety, Stress, & Coping*, 21: 173-184.

Kahn, W. A. 1990. Psychological conditions of personal engagement and disengagement at work. *Academy of Management Journal*, 33: 692-724.

Kahneman, D., Fredrickson, B. L., Schreiber, C. A., & Redelmeier, D. A. 1993. When more pain is preferred to less: Adding a better end. *Psychological Science*, 4: 401-405.

Karasek, R. A., Jr. 1979. Job demands, job decision latitude, and mental strain: Implications for job redesign. *Administrative Science Quarterly*, 24: 285-308.

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I. 2010. Control and interference in task switching—A review. *Psychological Bulletin*, 136: 849-874.

Kühnel, J., Sonnentag, S., & Bledow, R. 2012. Resources and time pressure as day-level antecedents of work engagement. *Journal of Occupational and Organizational Psychology*, 85: 181-198.

Lazarus, R. S., & Folkman, S. 1984. *Stress, appraisal, and coping*. New York, NY: Springer.

Le, H., Schmidt, F. L., & Putka, D. J. 2009. The multifaceted nature of measurement artifacts and its implications for estimating construct-level relationships. *Organizational Research Methods*, 12: 165-200.

LePine, J. A., Podsakoff, N. P., & LePine, M. A. 2005. A meta-analytic test of the challenge stressor-hindrance stressor framework: An explanation for inconsistent relationships among stressors and performance. *Academy of Management Journal*, 48: 764-775.

LePine, M. A., Zhang, Y., Crawford, E. R., & Rich, B. L. 2016. Turning their pain to gain: Charismatic leader influence on follower stress appraisal and job performance. *Academy of Management Journal*, 59: 1036-1059.

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.

McCormick, B. W., Reeves, C. J., Downes, P. E., Li, N., & Ilies, R. 2020. Scientific contributions of within-person research in management: Making the juice worth the squeeze. *Journal of Management*, 46: 321-350.

Miceli, M., & Castelfranchi, C. 2005. Anxiety as an “epistemic” emotion: An uncertainty theory of anxiety. *Anxiety, Stress, and Coping*, 18: 291-319.

Nahrgang, J. D., Morgeson, F. P., & Hofmann, D. A. 2011. Safety at work: A meta-analytic investigation of the link between job demands, job resources, burnout, engagement, and safety outcomes. *Journal of Applied Psychology*, 96: 71-94.

Parker, S. K., Morgeson, F. P., & Johns, G. 2017. One hundred years of work design research: Looking back and looking forward. *Journal of Applied Psychology*, 102: 403-420.

Pearsall, M. J., Ellis, A. P., & Stein, J. H. 2009. Coping with challenge and hindrance stressors in teams: Behavioral, cognitive, and affective outcomes. *Organizational Behavior and Human Decision Processes*, 109: 18-28.

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.

Podsakoff, N. P., Spoelma, T. M., Chawla, N., & Gabriel, A. S. 2019. What predicts within-person variance in applied psychology constructs? An empirical examination. *Journal of Applied Psychology*, 104: 727-754.

Rich, B. L., Lepine, J. A., & Crawford, E. R. 2010. Job engagement: Antecedents and effects on job performance. *Academy of Management Journal*, 53: 617-635.

Rodell, J. B., & Judge, T. A. 2009. Can “good” stressors spark “bad” behaviors? The mediating role of emotions in links of challenge and hindrance stressors with citizenship and counterproductive behaviors. *Journal of Applied Psychology*, 94: 1438-1451.

Rosseel, Y. 2012. lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48: 1-36.

Scheck, C. L., Kinicki, A. J., & Davy, J. A. 1995. A longitudinal study of a multivariate model of the stress process using structural equations modeling. *Human Relations*, 48: 1481-1510.

Schmidt, F. L., & Hunter, J. E. 2015. *Methods of meta-analysis: Correcting error and bias in research findings*. Thousand Oaks, CA: Sage.

Schneider, W., & Shiffrin, R. M. 1977. Controlled and automatic human information processing: Detection, search, and attention. *Psychological Review*, 84: 1-66.

Shaw, J. B., & Weekely, J. A. 1985. The effects of objective work-load variations of psychological strain and post-work-load performance. *Journal of Management*, 11: 87-98.

Spector, P. E., Chen, P. Y., & O’Connell, B. J. 2000. A longitudinal study of relations between job stressors and job strains while controlling for prior negative affectivity and strains. *Journal of Applied Psychology*, 85: 211-218.

Steed, L. B., Swider, B. W., Keem, S., & Liu, J. T. 2019. Leaving work at work: A meta-analysis on employee recovery from work. *Journal of Management*. Advance online publication. doi:10.1177/0149206319864153

Tims, M., & Bakker, A. B. 2010. Job crafting: Towards a new model of individual job redesign. *SA Journal of Industrial Psychology*, 36: 1-9.

Tims, M., Bakker, A. B., & Derks, D. 2012. Development and validation of the job crafting scale. *Journal of Vocational Behavior*, 80: 173-186.

Tims, M., Bakker, A. B., & Derks, D. 2013. The impact of job crafting on job demands, job resources, and well-being. *Journal of Occupational Health Psychology*, 18: 230-240.

Viswesvaran, C., & Ones, D. S. 1995. Theory testing: Combining psychometric meta-analysis and structural equations modeling. *Personnel Psychology*, 48: 865-885.

Wallace, J. C., Edwards, B. D., Arnold, T., Frazier, M. L., & Finch, D. M. 2009. Work stressors, role-based performance, and the moderating influence of organizational support. *Journal of Applied Psychology*, 94: 254-262.

Webster, J. R., Beehr, T. A., & Love, K. 2011. Extending the challenge-hindrance model of occupational stress: The role of appraisal. *Journal of Vocational Behavior*, 79: 505-516.

Yu, J. J., Downes, P. E., Carter, K. M., & O’Boyle, E. H. 2016. The problem of effect size heterogeneity in meta-analytic structural equation modeling. *Journal of Applied Psychology*, 101: 1457-1473.