TILBURG ◆ ◆ UNIVERSITY

Tilburg University

The application of the multilevel paradigm in human resource management outcomes research

Peccei, R.E.; van de Voorde, F.C.

The Application of the Multilevel Paradigm in Human Resource Management–Outcomes Research: Taking Stock and Going Forward

Riccardo Peccei
King's College London

Karina Van De Voorde
Tilburg University

The authors provide a systematic evaluation of the emerging multilevel paradigm in human resource management (HRM) research focusing, in particular, on multilevel mediation analyses of the HRM-outcomes relationship. They first distinguish different types of multilevel models and then identify a number of best practice theoretical and methodological criteria for the conduct of multilevel research derived from the literature. These criteria are used to analyze and evaluate all multilevel mediation studies (N = 46) of the HRM-outcomes relationship published in 11 core management and HRM journals in the 2000 to 2016 period. The results suggest that progress in the application of the multilevel paradigm has been slow and uneven and that the new paradigm has not as yet taken firm root in the field of HRM. On the basis of this evaluation, the authors identify key areas for improvement in the application of the multilevel approach to the analysis of the HRM-outcomes relationship. In the process, they highlight important ways in which the further development of a multilevel perspective can enhance both strategic and employee-centered HRM research by contributing to a fuller substantive understanding of the processes and mechanisms through which HRM systems affect outcomes of interest at different levels of analysis.

Keywords: *strategic HRM; HRM–outcomes research; multilevel paradigm; multilevel mediation*

Corresponding author: Karina Van De Voorde, Department of Human Resource Studies, Tilburg University, P. O. Box 90153, Tilburg 5000 LE, Netherlands.

E-mail: f.c.v.d.voorde@uvt.nl

Over the past quarter of a century, there has been a rapid growth in multilevel theory and research in management designed to bridge the gap between micro (e.g., individual) and macro (e.g., organizational) phenomena and levels of analysis (Aguinis, Boyd, Pierce, & Short, 2011). The move from a single-level to a multilevel perspective has major implications for theory, measurement, and analysis (Kozlowski & Klein, 2000) and is therefore commonly thought to represent a fundamental paradigm shift in management research (Hitt, Beamish, Jackson, & Mathieu, 2007; Mathieu & Chen, 2010). The spread of the multilevel paradigm, however, has been uneven across different subdomains of management. In particular, the field of human resource management (HRM) has lagged behind other areas of management scholarship in this respect, with the dominant paradigm in HRM still being one that is rooted in a single-level perspective typically focusing on the individual or organizational level (Molloy, Ployhart, & Wright, 2011).

Despite the continuing dominance of single-level studies, however, there is little doubt that the multilevel paradigm is gaining in importance also in the area of HRM. Over the past decade or so, in response to growing calls for more multilevel theory and research in HRM (K. Jiang, Takeuchi, & Lepak, 2013; Lengnick-Hall, Lengnick-Hall, Andrade, & Drake, 2009; Lepak, Jiang, Han, Castellano, & Hu, 2012; Peccei, Van De Voorde, & Van Veldhoven, 2013; Wright & Boswell, 2002), an increasing number of studies have adopted a multilevel approach to explore how the influence of HRM unfolds across levels in organizations (K. Jiang et al., 2013). Particularly important in this respect, and the central focus of the present analysis, are multilevel mediation studies of the HRM-outcomes relationship designed to examine the mechanisms through which multiple human resource (HR) practices (i.e., HRM bundles and systems; Wright & Boswell, 2002) affect different dimensions of organizational performance, as well as a range of key employee-related outcomes. Although at the very heart of HRM scholarship, the analysis of the HRM systems-outcomes relationship, or HRM-outcomes relationship for short, and of the employee-related mediating mechanisms involved, has been dominated by a single-level perspective (Guest, 2011; Jackson, Schuler, & Jiang, 2014; K. Jiang, Lepak, Hu, & Baer, 2012). The adoption of a multilevel approach, therefore, has the potential to make a significant contribution to the further development of this core area of HRM research. As such, it is now essential, we believe, to stand back and critically review progress with a view to identifying possible areas for improvement in the application of the multilevel paradigm that can help to advance HRM-outcomes research.

Such an exercise will, we hope, help to address an important gap in the HRM literature. A number of recent theoretical, methodological, and empirical reviews of HRM-outcomes research have touched on multilevel issues (Gerhart, 2013; Jackson et al., 2014; K. Jiang et al., 2013; Lengnick-Hall et al., 2009; Lepak & Boswell, 2012; Lepak et al., 2012; Peccei et al., 2013). With the exception of the K. Jiang et al. (2013) study, however, none of these reviews focused specifically on multilevel analysis and the application of the new paradigm in extant HRM-outcomes research. But while Jiang and his colleagues proposed an elaborate multilevel model for the analysis of strategic HRM (SHRM), they provided only a fairly limited overview of extant multilevel research in the area. In the present study, we draw on the Jiang et al. review of multilevel research but extend and develop their analysis in a number of important ways. First, we extend their review by bringing the analysis up to date and covering a much larger and more varied set of studies in the area. Second, when critically

examining extant multilevel mediation studies of the HRM-outcomes relationship, we pay systematic attention not only to SHRM studies concerned with the relationship between HR systems and various aspects of organizational performance but also to studies that focus explicitly on the impact of HR systems on employee-related outcomes, including key attitudes, behaviors, and dimensions of their well-being at work. Third, we examine the various studies involved in greater detail and explicitly assess them against a number of core theoretical and methodological criteria for the analysis of multilevel mediation derived from the literature. Finally, on the basis of this analysis, we identify a number of key multilevel issues that we suggest are central to a deeper understanding of the HRM-outcomes relationship and that therefore require systematic attention in future research in this area.

More generally, we have two broad aims in this study. First is to look back and consider what has and what has not been achieved in terms of the application of the new multilevel paradigm to the analysis of the HRM-outcomes relationship. To this end, we first identify a number of best practice theoretical and methodological criteria for the conduct of multilevel research derived from the literature (Chan, 1998; Chen, Mathieu, & Bliese, 2005; Croon, van Veldhoven, Peccei, & Wood, 2014; Dansereau & Yammarino, 2000; Fulmer & Ostroff, 2016; Kozlowski & Klein, 2000; Mathieu & Taylor, 2007; Preacher, Zyphur, & Zhang, 2010; Shen, in press; Tay, Woo, & Vermunt, 2014). We then use these criteria to evaluate extant multilevel mediation studies of the HRM-outcomes relationship with a view to providing a systematic critical overview of the current state of development, application, and contribution of the multilevel paradigm to HRM-outcomes research. The second aim is to look ahead and consider how a multilevel perspective can contribute to the further development of HRM-outcomes research. Specifically, on the basis of our evaluation of extant multilevel studies in the area, we identify key areas for improvement in the application of the multilevel approach to the analysis of the HRM-outcomes relationship. In the process, we highlight important ways in which the further development and explication of a multilevel approach can contribute to a fuller substantive understanding of the processes and mechanism through which HR systems affect key outcomes of interest at different levels of analysis.

The Emerging Multilevel Paradigm in HRM-Outcomes Research

Single-level versus multilevel models. When examining the development of the multilevel paradigm in HRM, it is useful to start by distinguishing between two broad strands of HRM-outcomes research depending on whether the main focus is on the link between HRM and organizational performance outcomes or on the link between HRM and employee-related outcomes. Specifically, the first and, historically, by far the most important and extensive stream of HRM-outcomes research is in the area of SHRM. The core focus of SHRM is on the relationship between HRM and organizational performance—on the effect that sets or bundles of HR practices have on various aspects of both proximal and distal unit-level performance (e.g., organizational turnover, productivity, service quality, and overall financial performance; Arthur, 1994; Becker & Huselid, 1998; Huselid, 1995). The second more recent but growing stream of HRM-outcomes research is more explicitly focused on employee outcomes. The primary concern of what for ease of presentation we refer to as employee-centered HRM (ECHRM) is not necessarily with the potential effect of bundles of HR practices on organizational performance as such. Rather, the core focus is on the impact that HR systems have on employee experiences, attitudes, and behaviors. The employee

outcomes of interest here include not only various dimensions of individual task and contextual performance (e.g., in-role and extrarole behaviors) but also, and increasingly, various aspects of employee psychological and health-related well-being (e.g., job satisfaction, positive affect, job stress, burnout; Godard, 2001; Guest, 1999, 2011).

A detailed discussion of these two main strands of HRM-outcomes research is beyond the scope of the present analysis. Interested readers are referred to a number of wide-ranging theoretical and empirical overviews of SHRM and ECHRM research published in recent years (Boselie, Dietz, & Boon, 2005; Guest, 2011; K. Jiang et al., 2013; Lengnick-Hall et al., 2009; Wall & Wood, 2005; Wright & Gardner, 2003), including major meta-analyses of the HRM-performance relationship (Combs, Liu, Hall, & Ketchen, 2006; K. Jiang et al., 2012; Subramony, 2009) as well as qualitative reviews of ECHRM research focusing specifically on employee well-being (Peccei et al., 2013; Van De Voorde, Paauwe, & Van Veldhoven, 2012). Here it is sufficient to note that HRM-outcomes research, whether in the SHRM or ECHRM tradition, has historically been dominated by single-level studies. Based on a macroperspective, SHRM studies have concentrated primarily on the unit level, focusing on the impact of HR practices on different aspects of performance at the level of the group, establishment, or organization as a whole (K. Jiang et al., 2013). ECHRM studies, on the other hand, have focused primarily on the individual level, linking employee attitudes and behaviors with individual perceptions and experiences of HR practices (Peccei et al., 2013). Whether at the individual or unit level, however, the dominant single-level approach to the analysis of the HRM-outcomes relationship has evolved over time both theoretically and empirically. In particular, as reflected in the so-called black box debate in both SHRM (Guest, 2011; Wright & Gardner, 2003) and ECHRM (Peccei & Van De Voorde, 2014), a key way in which standard single-level models in the area have been refined is through the development and testing of more elaborate models incorporating various types of employee-related mediating mechanisms, such as human capital, perceived organizational support, empowerment, and organizational climate. Importantly, however, these more elaborate mediation models are all still focused either at the individual or at the unit level of analysis and, consequently, remain rooted in a single-level paradigm.

In this context, multilevel mediation models of the HRM-outcomes relationship can be seen not only as a further elaboration and refinement of standard single-level models but also as an important departure from such models. The distinguishing feature of these multilevel mediation models is that, in contrast to single-level models, they incorporate independent (HRM), dependent (outcome), and mediating constructs that are explicitly treated as residing at potentially different levels of analysis (e.g., individual and organizational). More generally, central to these multilevel models and to the multilevel perspective more broadly is the idea that "organizational entities reside in nested arrangements" (Mathieu & Chen, 2010: 610). For example, individuals are nested in organizational subunits (e.g., work groups) that are, in turn, nested in organizations.

There are two main implications of such nesting. First, from a theoretical perspective, nested data structures mean that in order to understand phenomena at a given level of analysis, it is often not sufficient just to focus on factors operating at that same level. Rather, it is necessary also to understand so-called cross-level relationships and to take into account factors at higher or lower levels that may influence focal constructs of interest residing at a different level (Raudenbush & Bryk, 2002). Micro-ECHRM research that focuses at the individual level, for example, tends to neglect contextual factors, such as unit-level HR

practices, that can influence and constrain the experiences, attitudes, and/or behaviors of individuals within those units. Equally, though, macro-SHRM research that focuses at the unit level may need to consider not only how unit-level HR practices affect the attitudes and behaviors of employees within these units but also how the attitudes and behaviors of employees within the units may, in turn, affect aspects of the unit's performance (Kozlowski & Klein, 2000).

Second, and from a more purely methodological standpoint, because of the hierarchical nature of nested data, observations based on clustered data of these kinds may not be independent, resulting in biased parameter estimates of interest (Kreft & de Leeuw, 1998). Until relatively recently, the main strategy for dealing with cross-level effects in hierarchically nested data was to opt for either aggregation or disaggregation. Aggregation involves aggregating the lower-level variables in a model (e.g., individual job satisfaction) to a higher unit level of analysis and then using the aggregated variable (i.e., mean unit-level job satisfaction), together with other higher unit-level variables in subsequent analyses (Hofmann, Griffin, & Gavin, 2000). Disaggregation, on the other hand, involves assigning to each lower-level entity in a nested structure (e.g., individuals within organizations) the corresponding score of a higher-order unit-level variable (e.g., the HRM score of the organization in which the individual is employed; Hofmann et al., 2000). For a variety of reasons, both strategies are highly problematic (Kreft & de Leeuw, 1998; Raudenbush & Bryk, 2002). The preferred alternative to test more complex mediational relationships involving explanatory HRM, mediating, and outcome variables residing at different levels of analysis is to use a multilevel approach explicitly designed to capture cross-level effects in hierarchically nested data (Snijders & Bosker, 1999).

Types of HRM multilevel mediation models. Different types of multilevel mediation models of the HRM-outcomes relationship have been proposed and tested in the SHRM and ECHRM literature. Based on the work of Kozlowski and Klein (2000), Mathieu and Taylor (2007), and Preacher et al. (2010), these models can be broadly classified into different basic types depending on the direction of operation ("upward" and/or "downward") of the key cross-level mediational processes involved. By convention, variables located and conceptualized at the lower (e.g., individual) and higher (e.g., unit) level in multilevel models are referred to as Level 1 and Level 2 variables, respectively. We start by considering the models used in SHRM research where the focus is on some form of organizational performance as the key outcome of interest.

The first broad type of multilevel mediation model used in extant SHRM-outcomes studies is what Croon et al. (2014) refer to as a 2-1-2 mixed mediation model. This type of model involves both downward and upward effects, such as the downward influence of a Level 2 HRM variable (e.g., a system of HR practices at unit level) on a Level 1 variable (e.g., individual employee job satisfaction) that, in turn, is hypothesized to have an upward effect on a Level 2 outcome variable (e.g., the quality of customer service in the unit). In practice, as we discuss more fully below, 2-1-2 models can vary considerably in terms of their complexity depending, for example, on the number and combination of Level 2 and Level 1 variables involved and the complexity of the mediation paths that are hypothesized. Irrespective of their complexity, however, underpinning 2-1-2 models is the idea, central to an important line of thinking in SHRM, that the way bundles of HR practices ultimately affect organizational performance is through people (Becker & Huselid, 1998; K. Jiang et al., 2013; Pfeffer, 1998;

Wright & Boswell, 2002). Specifically, based on a variety of theoretical arguments linked, for example, to the resource-based view (Den Hartog, Boon, Verburg, & Croon, 2013; Gong, Law, Chang, & Xin, 2009), human capital theory (Takeuchi, Lepak, Wang, & Takeuchi, 2007), the behavioral perspective in HRM (Messersmith, Patel, Lepak, & Gould-Williams, 2011), and the ability, motivation, and opportunity (AMO) framework (Raineri, in press), SHRM-related 2-1-2 type models focus on the Level 2 HRM–firm performance relationship and posit Level 1 employee attitudes and behaviors as key mediators of this relationship.

Importantly, however, as part of SHRM research, we also identified a type of mediation model that, while incorporating elements of a multilevel perspective, remained rooted in a single-level approach (Kozlowski & Klein, 2000). The distinguishing feature of what might be termed proto-multilevel (proto-ML) models of this kind is that although they are based on nested data, they adopt a simple aggregate approach of the type described above to the analysis of multilevel data. That is to say, any lower Level 1 independent, mediator, or dependent variables in the analysis are simply aggregated to the unit level (Level 2) and then used as aggregate variables together with other higher-order unit-level variables in what can essentially be thought of as single-level 2-2-2 type models. These aggregate-level 2-2-2 models are then tested using standard analytic techniques, rather than multilevel procedures. Unlike in standard single-level 2-2-2 studies, however, in proto-ML 2-2-2 studies, an explicit attempt is made to justify, theoretically and/or empirically, the aggregation of Level 1 variables to the unit level by reference, for example, to a variety of key aggregation statistics (e.g., r_{wg} and intraclass correlation coefficients ICC1 and ICC2). Proto-ML 2-2-2 studies, in other words, combine a single-level approach with some elements of a multilevel perspective. It should be noted that, in practice, depending on a number of factors, such as the number of Level 2 units involved, the size of Level 1 samples, and the extent of within- and between-group variance in nested data, the results obtained using the simple aggregation procedures of proto-ML analyses can be similar to those obtained using a more explicit multilevel approach characteristic of 2-1-2 models (Preacher et al., 2010). Nevertheless, the failure properly to take account of nesting in hierarchically structured data, as well as to distinguish within- from between-group variance in proto-ML models, makes the use of these models potentially more problematic (Croon et al., 2014; Preacher et al., 2010). This, together with the fact that in a number of key respects, proto-ML models remain rooted in a single-level paradigm, sets these models apart from 2-1-2 multilevel models and, therefore, justifies treating them as a separate category in the present analysis.

Apart from the mixed mediation and proto-ML models identified above, one other type of multilevel mediation model appears to have been used in extant SHRM research. This is an upward (1-1-2) mediation model used in a recent study that examined the relationship between employee perceptions of HR practices at Level 1 and unit performance (appropriateness of care) at Level 2, as mediated by individuals' perceptions of proactivity climate at Level 1 (Dello Russo, Mascia, & Morandi, in press).

In contrast to SHRM research, the main type of multilevel mediation model employed in extant ECHRM-outcomes research is a downward mediation model. Downward mediation examines the effect of a Level 2 HRM variable (e.g., a system of HR practices at unit level) on a Level 1 variable (e.g., individual employee customer-oriented behavior) as mediated either by another Level 1 variable (e.g., individual employee job satisfaction) or by another Level 2 variable (e.g., organizational climate for customer service). In other words, there are two main variants of downward mediation models that can be distinguished depending on

whether the mediator of interest is at Level 1 (i.e., a 2-1-1 model) or at Level 2 (i.e., a 2-2-1 model; Mathieu & Taylor, 2007). As with 2-1-2 mixed mediation models, downward mediation models, whether of the 2-1-1 or the 2-2-1 variety, can differ considerably not only in their complexity but also in terms of the specific theories that underpin them. Depending on the specific focus of the model, the theories involved range, for example, from social exchange and self-determination theory (Kehoe & Wright, 2013; Shin, Jeong, & Bae, in press), to organizational justice theory (Kroon, Van de Voorde, & Van Veldhoven, 2009; Wu & Chaturvedi, 2009), to empowerment and climate theory (Oppenauer & Van De Voorde, in press; Takeuchi, Chen, & Lepak, 2009). Below, we use the main types of mixed, upward, and downward models identified above, including the proto-ML models, to classify extant multilevel mediation studies of the HRM-outcomes relationship.

Framework for the Evaluation of Multilevel Mediation Studies

Central to any systematic assessment of the application and development of the new multilevel paradigm in HRM-outcomes research is the identification of core theoretical and methodological criteria in terms of which to evaluate progress in this area. To this end, we start by presenting a set of best practice criteria for the analysis of cross-level mediation and the conduct of multilevel research more generally against which extant multilevel studies of the HRM-outcomes relationship can be evaluated. The framework we used for our analysis covers both conceptual/theoretical and analytic/methodological criteria. The criteria themselves were derived from the literature drawing, in particular, on the general multilevel work of Chan (1998), Kozlowski and Klein (2000), Dansereau and Yammarino (2000), Chen, Mathieu, and Bliese (2005), Croon et al. (2014), Tay et al. (2014), and Fulmer and Ostroff (2016), as well as on the more specific best practice recommendations for the analysis of multilevel mediation of Mathieu and Taylor (2007), Preacher et al. (2010), and Shen (in press). Before considering the proposed best practice criteria in greater detail, however, we believe a number of general points about the present framework should be noted.

First, it is important to bear in mind that the best practice criteria outlined below are necessarily time dependent and relative and should be viewed accordingly. Because of the development of relevant theoretical and methodological knowledge, for example, what might be considered to be best practice in multilevel research at one point in time may not necessarily be considered to be equally optimal at a later date. Moreover, for a variety of reasons, best practice recommendations may take time to diffuse and become widely accepted across a community of scholars (Hitt et al., 2007; Kuhn, 1962). In addition, it is important to recognize that some areas of multilevel analysis, such as the analysis of upward 1-2 links, are still relatively underdeveloped (Shen, in press). Knowledge in these areas is still in relative flux so that best practice recommendations in these areas need to be treated with some caution.

Second, the criteria in the framework are interdependent in that conceptual/theoretical and analytic/methodological choices in multilevel analysis are closely interrelated. The type of multilevel model that is adopted and the way in which core constructs are conceptualized in such models, for example, have important implications not only for theory development but also for the type of analytical strategies and methodological approaches that are used to test cross-level hypotheses of interest. As we discuss more fully below, a key overall criterion for the evaluation of multilevel studies in fact is the extent to which there is an appropriate alignment between conceptual/theoretical and analytic/methodological elements of the research.

As highlighted by Wallace, Edwards, Paul, Burke, Christian, and Eissa, such alignment is important since "if there is a mismatch between the level of theory, measurement, and/or statistical analysis, empirical results may not be interpretable with regard to theoretical or methodological linkages among the constructs" (2016: 840).

The third, and related, general point to note is more technical in nature but has important implications for the theoretical understanding and analysis of multilevel mediation. The point in question concerns the importance, long noted in the general multilevel literature, of disentangling between- and within-group effects in multilevel analysis (Dansereau & Yammarino, 2000; Hofmann, 1997; Kreft & de Leeuw, 1998). Specifically, as emphasized by Preacher et al. (2010), central to the analysis of cross-level mediation effects is the recognition that while Level 2 constructs have only between components of variance, Level 1 variables typically have both between and within components of variance, with the between components corresponding to the mean group-level scores of the observed Level 1 variables. Therefore, for "any mediation model involving at least one level-2 variable, the indirect effect can exist only at the between level" (Preacher et al., 2010: 213). In a 2-1-1 design, for example, the indirect effect of an HR system (a Level 2 independent variable) on employee job performance (a Level 1 outcome variable) through individual job satisfaction (a Level 1 mediator) may function only through the between-group variance in the mediator and outcome variables. More generally, as we discuss below, this insight, based on a partitioning perspective (see also the within and between analytical approach; Dansereau & Yammarino, 2000), has important implications not only for construct conceptualization but also for the nature of theoretical explanations and for the analytic procedures used in multilevel mediation studies.

With these general considerations in mind, we now turn to the specific best practice criteria we used to evaluate the various multilevel mediation studies of the HRM-outcomes relationship we examined. The main key best practices involved are summarized in Table 1.

Conceptual and theoretical criteria. The first set of dimensions of the framework (Criteria 1a–1e) relate to the conceptualization of the Level 2 and Level 1 constructs included in multilevel mediation models. As emphasized by Kozlowski and Klein, "The first and foremost task in crafting a multi-level theory or study is to define, justify, and explain the level of each focal construct that constitutes the theoretical system" (2000: 27). In practice, this includes a number of specific steps, the first of which is to provide a clear definition of all main Level 2 and Level 1 constructs in the hypothesized model (Criterion 1a). Importantly, in line with the partitioning perspective emphasized above, this includes providing explicit definitions and conceptualizations of all main variables in the analysis. This covers not only directly assessed Level 2 and Level 1 model variables but also what, following Preacher et al. (2010), we refer to as Level 2 "derived between" constructs in multilevel mediation models. The latter constructs refer to the between components corresponding to the mean group-level scores of the observed Level 1 constructs, for example, mean group satisfaction scores, representing the between-level component of a Level 1 job satisfaction variable.

Having defined all relevant Level 2 constructs in a multilevel model, the next important step, in terms of conceptualization, is to specify the precise nature of all the Level 2 directly assessed and "derived between" constructs in the analysis (Criterion 1b). In particular, drawing on the work of Chan (1998) and of Kozlowski and Klein (2000), our framework highlights the importance of distinguishing between four main types of Level 2 constructs commonly used in multilevel research. These include (a) global, (b) aggregate rated or

Table 1
Key Best Practices for Conducting Multilevel Mediation Studies

1. Conceptualization of constructs
   a. Explicitly define Level 2 model and derived constructs, and Level 1 constructs
   b. Specify the overall nature of all Level 2 model constructs (i.e., whether global, aggregate rated/additive, aggregate consensus, or aggregate compilational)
   c. Specify the type of composition of aggregate consensus Level 2 model constructs (i.e., direct consensus vs. referent shift)
   d. Provide an explicit explanation of how aggregate consensus Level 2 constructs emerge/manifest at Level 2
   e. Theorize explicitly about the extent to which Level 2 aggregate compositional constructs have similar meanings and properties as their Level 1 counterparts (cross-level psychometric isomorphism)
2. Theoretical justification
   a. Provide explicit theoretical justification of any 2-1, 1-2, 1-1, or 2-2 links in the model
   b. Theorize explicitly about the extent to which relationships between constructs at Level 2 are generalizable at Level 1 (cross-level homology)
3. Aggregation
   a. Provide ICC1s for all model constructs and all Level 2 aggregate derived constructs to ensure there is sufficient between-unit variation to justify multilevel analysis
   b. Provide ICC2s and r_{wg} for all aggregate consensus Level 2 constructs to ensure there is sufficient within-unit agreement to justify aggregation
4. Model testing
   a. Test the hypothesized (cross-level) links simultaneously in an integrated analytic framework
5. Modeling of within- and between-group effects
   a. Examine separately the within-group and between-group level effects
   b. Report the separate R^2 values for within- and between-group analyses

Note: ICC = intraclass correlation coefficient.

additive, (c) aggregate consensus, and (d) aggregate compilational supraindividual Level 2 constructs, respectively. The reason why it is important to distinguish between the four types is that the specific nature of the Level 2 constructs included in multilevel research has important theoretical and methodological implications for the analysis. For example, as we discuss below, different types of Level 2 constructs require not only different theoretical explication but also the use of different analytical procedures.

A detailed discussion of the four types of supraindividual constructs is beyond the scope of the present study. Briefly, global Level 2 constructs involve an overall summary rating of a supraindividual phenomenon (e.g., unit-level ratings by a single manager of the organization's HR practices, or objective overall measures of unit-level productivity). In contrast, aggregate compositional Level 2 constructs involve the summation of individual-level scores on a particular variable (e.g., job satisfaction) to the supraindividual level (e.g., mean unit-level job satisfaction). In the case of aggregate compositional constructs, however, we distinguish between constructs based on what Chan (1998) refers to as additive versus consensus forms of composition. The main difference between these two kinds of compositional constructs is that in the case of additive forms of composition, the conceptual validity of the aggregate construct does not necessarily depend on the existence of a high level of within-group agreement on the variable of interest. This is the case, for example, with mean unit-level manager ratings of HR practices based on the summation of the ratings of individual managers. In this case, the managers are merely used as informants with access to relevant information, and their ratings are simply averaged (Kozlowski & Klein, 2000). We therefore

refer to these constructs as *rated or additive* aggregate Level 2 compositional constructs. In the case of consensus-based forms of composition, on the other hand, "the meaning of the higher level construct is in the consensus among lower level units" (Chan, 1998: 237; e.g., the shared unit-level climate perceptions of members of the unit). As these shared properties of a unit emerge from employees' perceptions, affect, and behaviors originating at the individual level (Kozlowski & Klein, 2000), following Chan, we refer to these constructs as *consensus-based* aggregate Level 2 compositional constructs.

The last type of aggregate Level 2 construct included in the framework is a compilational one. These constructs are based on potentially more complex processes of emergence and aggregation than compositional constructs. Above all, unlike consensus-based compositional constructs, the meaning of compilational constructs is not in the agreement among lower units. Rather, as noted by Chan (1998) and by Kozlowski and Klein (2000), the meaning of higher-level compilational constructs is in the dispersion, variance, or other forms of patterning of scores among lower-level units. As such, higher-level compilational constructs are not captured or represented by the simple aggregation (e.g., averaging) of lower-level scores to the supraindividual level. Rather, they are typically represented by more complex Level 2 measures based, for example, on different patterns of dispersion or on different combinations or configurations of Level 1 scores. Hence, the partitioning perspective outlined above does not apply to the analysis of cross-level relationships that involve more complex types of compilational Level 2 variables. As we discuss more fully below, this has important implications for multilevel mediation theory and research.

Depending on the types of Level 2 constructs that are included in multilevel studies, three further best practice criteria come into play. Specifically, for consensus-based compositional constructs, it is important to identify the particular form of composition that is involved and specify whether the construct is based on direct consensus or referent-shift forms of aggregation (Chan, 1998; Wallace et al., 2016; Criterion 1c). In addition, for consensus-based compositional constructs, it is essential to articulate the processes by which these higher-level constructs form. In particular, as emphasized by Chan (1998) and by Kozlowski and Klein (2000), it is important to identify the compositional processes and mechanisms through which within-unit consensus and convergence in individual attitudes and behaviors might develop. The aim, drawing, for example, on social information processing theory or the attraction-selection-attrition (ASA) framework (Fulmer & Ostroff, 2016), is both to explain the emergence of the relevant Level 2 constructs from the corresponding Level 1 variables and to justify the aggregation of the Level 1 scores to the higher level of analysis (Criterion 1d). Note that the processes and mechanisms that underpin the emergence of more complex Level 2 compilational constructs also need explication in multilevel studies. As we discuss more fully below, however, none of the multilevel HRM-outcomes studies we examined considered models that included compilational Level 2 constructs. Hence, in our framework, we highlight only Criterion 1d that relates to construct explication with respect to consensus-based compositional constructs.

Finally, in the case of compositional constructs, it is also important to explore cross-level psychometric isomorphism issues and explicitly consider whether Level 2 compositional constructs have similar meanings and properties as their Level 1 counterparts. The psychometric equivalence of multilevel constructs has received increasing attention not only in the multilevel literature (Tay et al., 2014) but also in specific areas of management research

(Dyer, Hanges, & Hall, 2005). As noted by Tay et al. (2014), in fact, an examination of the psychometric equivalence of parallel constructs at different levels of analysis is important not only for the exploration of homology relations across levels but also for the development of multilevel theory and research more generally. Hence, we included psychometric isomorphism as one of the evaluation criteria in our best practice framework (Criterion 1e).

The next set of criteria in the framework relate to a number of key theoretical issues involved in multilevel mediation research (Criteria 2a and 2b). Specifically, central to the development and testing of multilevel mediation models is the need theoretically to justify all the main hypothesized cross-level (2-1; 1-2) and single-level (1-1; 2-2) links in the model (Mathieu & Taylor, 2007; Shen, in press; Criterion 2a). Particularly important in this respect is the need to explain all proposed cross-level links between independent, mediator, and outcome variables by reference to appropriate microtheories and/or macrotheories capable of accounting for both downward and upward effects in multilevel mediation models of the HRM-outcomes relationship (K. Jiang et al., 2013; Peccei et al., 2013). The specific theories that might be invoked here include, for example, social exchange theory, psychological climate theory, the AMO framework, aspects of human capital theory, the job demands–resources model, or resource-based view arguments (Jackson et al., 2014; K. Jiang et al., 2013; Peccei et al., 2013).

The last theory-related criterion in the framework concerns the theorization of cross-level homology in multilevel mediation models (Criterion 2b). In multilevel research, homology refers to the extent to which the relationship among constructs at one level of analysis is more or less comparable to the relationship between corresponding or functionally equivalent constructs at a different level of analysis (Kozlowski & Klein, 2000). Although homology issues have been extensively debated in the multilevel literature (Chen, Bliese, & Mathieu, 2005), they have received only very limited attention in HRM research (K. Jiang et al., 2013). Because of the potential importance of cross-level homology in multilevel mediation models of the HRM-outcomes relationship, however, we included a homology criterion in our framework.

Measurement and analysis criteria. Analytically, a key first step that best practice suggests is central to the conduct of multilevel mediation research is to determine the proportion of the variance in the relevant dependent, mediation, and outcome variables of a multilevel model that resides at different levels and check that there is sufficient variance at Level 2 to justify the use of a multilevel approach in the first place (Shen, in press). In the absence of such information, it is difficult to determine the potential contribution a multilevel approach can make to our understanding of the mediating mechanisms in the HRM-outcomes relationship, which necessarily operate at the between-unit level (Preacher et al., 2010). Hence, the first methodological criterion included in our framework asks whether relevant aggregation statistics and within- and between-group variance statistics are estimated for all relevant study variables, including both main model variables and "derived between" Level 2 constructs (Criterion 3). In particular, this includes ICC1 values for all Level 2 aggregate model variables and all Level 2 derived between constructs in the analysis (Criterion 3a). It also includes ICC2 (Bliese, 2000) and r_{wg} (James, Demaree, & Wolf, 1984) or average deviation (Burke, Finkelstein, & Dusig, 1999) values for all Level 2 aggregate consensus-based constructs (Shen, in press; Criterion 3b).

Having established the appropriateness of using a multilevel approach, the next main analysis-related criterion relates to model testing (Criterion 4a). Specifically, this criterion asks whether hypothesized model relationships, including in particular cross-level mediation links, are tested separately or simultaneously in an integrated analytic framework. This criterion can best be understood in relation to a more general point about the analysis of cross-level links in 2-1-2 mediation models. As noted by a number of researchers (Croon et al., 2014; Shen, in press), the analytic and statistical techniques needed for analyzing bottom-up 1-2 relationships with multilevel data are far less well-established and understood than those required to examine top-down 2-1 effects. To the extent that bottom-up 1-2 effects are examined, however, the tendency is to aggregate the Level 1 variable to the unit level and then conduct a separate analysis using standard multiple regression to examine the relationship between the new aggregated variable and the relevant Level 2 outcome (Aryee, Walumbwa, Seidu, & Otaye, 2012). An alternative to using this more disjointed type of approach to the analysis of 1-2 relationships is to use analytic procedures, such as multilevel structural equation modeling (MSEM), that can accommodate dependent variables measured at Level 2, thereby allowing for a more direct and effective analysis of 1-2 relationships as part of more complex multilevel mediation models (Muthén, 2007; Preacher et al., 2010; Shen, in press). It should be noted, however, that there is still debate in the methodological literature about the added value and robustness of some of these analytic procedures for the analysis of 1-2 links in multilevel models and that MSEM procedures in this area are still evolving (Croon et al., 2014). The overall aim, nevertheless, is to use a more integrated approach to the simultaneous testing of cross-level relationships in complex downward, upward, and full-mediation multilevel models of the HRM-outcomes relationship, hence accounting for the inclusion of this criterion in our framework.

The final criterion in the framework relates to the appropriate modeling and testing of within- and between-group effect (Criterion 5). A clear implication of the partitioning perspective is that when testing cross-level effects in multilevel mediation models, it is crucial properly to partition the variance in Level 1 variables into within- and between-group components. This is in order to separate within- from between-group level effects to arrive at unbiased estimates of relevant between effects in the data. Importantly, however, as noted by Preacher et al. (2010), the multilevel modeling (MLM) analysis procedures traditionally used for testing multilevel models do not enable researchers to distinguish within effects from between effects. MSEM procedures, on the other hand, enable researchers to model and test within and between effects separately and, therefore, are generally to be preferred to standard MLM approaches to multilevel mediation (Preacher et al., 2010; Shen, in press; Criterion 5a). As part of a good practice approach, and consistent with the partitioning perspective, this criterion also asks whether separate effect sizes (R^2 values) for within- and between-group analyses are reported (Criterion 5b).

Method

Study Search and Selection

A systematic literature search in 11 respected journals for articles dealing with the effects of HRM was conducted to identify relevant multilevel studies. These journals included *Academy of Management Journal*, *Human Relations*, *Human Resource Management*, *Human*

Resource Management Journal, *The International Journal of Human Resource Management*, *Journal of Applied Psychology*, *Journal of Management*, *Journal of Management Studies*, *Journal of Organizational Behavior*, *Personnel Psychology*, and *Personnel Review*. We selected these journals because they are seen as top journals in the general management and organizational behavior field (e.g., *Academy of Management Journal*, *Journal of Applied Psychology*) or because they represent dedicated HRM journals (e.g., *Human Resource Management Journal*). Since the first multilevel framework about HRM was developed in 2000 by Ostroff and Bowen and the need for more multilevel research in HRM was highlighted in 2002 by Wright and Boswell, we chose 2000 as the starting date for our literature search and reviewed the articles published between 2000 and May 2016 inclusive.

For a study to be included in our review, the study had to meet all of the following three criteria. First, following recent attempts in SHRM and ECHRM research to understand the processes through which HR practices influence outcomes, it had to consider the mediating role of employee perceptions, attitudes, or behaviors in the relationship between HRM and employee and/or organizational outcomes. Studies that explore moderation only and not also mediation (e.g., S. Chang, Jia, Takeuchi, & Cai, 2014) or studies that do not explicitly explore mediation (e.g., Wright, Gardner, Moynihan, & Allen, 2005) are not included. Second, the study had to test empirically hypothesized relationships between constructs including at least one unit-level construct composed of aggregate individual-level data (proto-ML models) or test hypothesized relationships between constructs at different levels of analysis (mixed, upward, and downward models). In other words, studies testing hypothesized relationships between constructs at a single level of analysis (e.g., Jensen, Patel, & Messersmith, 2013) or for which it was impossible to determine the level of analysis applied were excluded. Third, following the main focus of SHRM, the study had to explore the effect of the use of a broad set of HR practices rather than the effect of a limited/focused set of HR practices or the effectiveness/evaluations of these practices. Hence, for example, studies examining age-related HRM (Bal, Kooij, & De Jong, 2013) or HR attributions (Nishii, Lepak, & Schneider, 2008) were not included in the analysis.

Multilevel Mediation Studies Covered

This search yielded a total of 46 studies/articles (28 SHRM and 18 ECHRM, which are denoted with an asterisk in the literature references). Summary information about the basic design of each of the studies included in the analysis is presented in the appendix in the online supplemental material. To facilitate the presentation of information, we distinguish between the three types of SHRM and one type of ECHRM mediation model used: mixed multilevel models, proto-ML models, downward multilevel models, and upward multilevel models. The majority of SHRM studies can be classified as proto-ML (18 out of 28 studies), in that they investigated the effect of HRM on performance at Level 2 by aggregating at least one Level 1 HRM, mediator, or performance variable to the unit level (see the appendix). Nine studies explored a mixed multilevel mediation model investigating both the downward influence of HRM at the unit level on a Level 1 mediator and an upward influence of a Level 1 mediator on a Level 2 performance variable. Only 1 study focused exclusively on upward mediation (Dello Russo et al., in press). The 18 ECHRM studies explored a downward mediation process whereby HRM at the unit level affects a Level 1 performance variable as mediated by either a Level 1 or a Level 2 variable (16 and 2 studies,

respectively). Regarding the complexity of the multilevel models used, most of the studies (28 out of 46) examined elaborate mediation models covering multiple employee-related mediators (such as human capital, employee relationships, organizational climate, employee perceptions of HRM, job characteristics, employee attitudes, employee health, and employee behaviors) and/or incorporated extended path-analytic chains. In addition, 13 studies investigated some form of moderated mediation.

Turning to the nature of the multilevel sample employed, the number of Level 1 units in each study ranged between 225 and 101,169 employees, and the number of Level 2 units in each study ranged between 11 and 2,001. In terms of the type of Level 2 units involved, 22 studies focused at the level of the organization, 21 at the level of business units or teams, and 3 at the level of the job group. Although 23 studies reported that the Level 2 units examined were nested in wider organizations, industries, or countries (denoted between brackets in the appendix), only 3 studies took a third level of nesting explicitly into account in the analysis (Hong, Liao, Raub, & Han, 2016; Kehoe & Wright, 2013; Liao, Toya, Lepak, & Hong, 2009).

The two authors of this paper reviewed each article for the five identified criteria (see Table 1). After discussing a small number of initial discrepancies in coding, coders were able to agree on all of the codes on the five criteria. To facilitate the presentation of information, here we report only the proportion of studies meeting the various evaluation criteria. This information is shown in Figures 1, 2, and 3.

Results

The first conceptual criterion we used to evaluate extant multilevel mediation studies of the HRM-outcomes relationship concerns the provision of explicit definitions and conceptualizations of all main variables in the research. Figure 1 shows that in all studies, the Level 1 and 2 model variables were explicitly defined (Criterion 1a). Of the 46 studies we examined, 8 included derived between constructs. Although in most of these studies the analytical reasons for partitioning are presented, none of the studies explicitly defined what the between-group (and within-group) scores mean conceptually (Criterion 1a). Turning to the specification of the nature of the Level 2 employee-rated constructs included in the studies (Criterion 1b), in 13 out of the possible 34 studies involved here, all the global, aggregate rated, and/or aggregate consensus Level 2 constructs covered were explicitly specified; in 3 studies, only some of the constructs involved were explicitly specified; and in 2 studies, none of the constructs were explicitly specified. The remaining studies included some form of aggregate compositional constructs, but the precise nature of these constructs (rated vs. consensus) remained implicit and had to be inferred from the way they were defined and/or from available descriptions of relevant emergence processes. Moreover, in the majority of the studies (14 out of 24) that included some form of consensus-based compositional constructs, the form of composition involved (direct consensus or referent shift) was not explicitly identified (Criterion 1c). However, in terms of the theoretical justification for the aggregation of these Level 2 consensus-based variables (Criterion 1d), in 21 out of the 24 studies involved, extensive justification was provided by reference to a variety of compositional processes and mechanisms likely to foster the development of within-unit consensus and convergence in employee perceptions, attitudes, and behaviors. Finally, even though (psychometric) isomorphism was mentioned in 6 studies, it was neither discussed in detail nor tested for in any of the studies we examined (Criterion 1e).
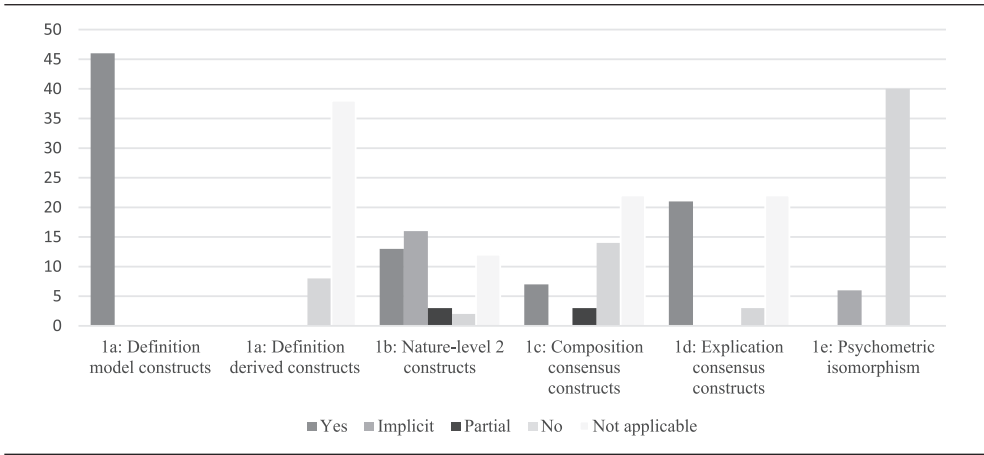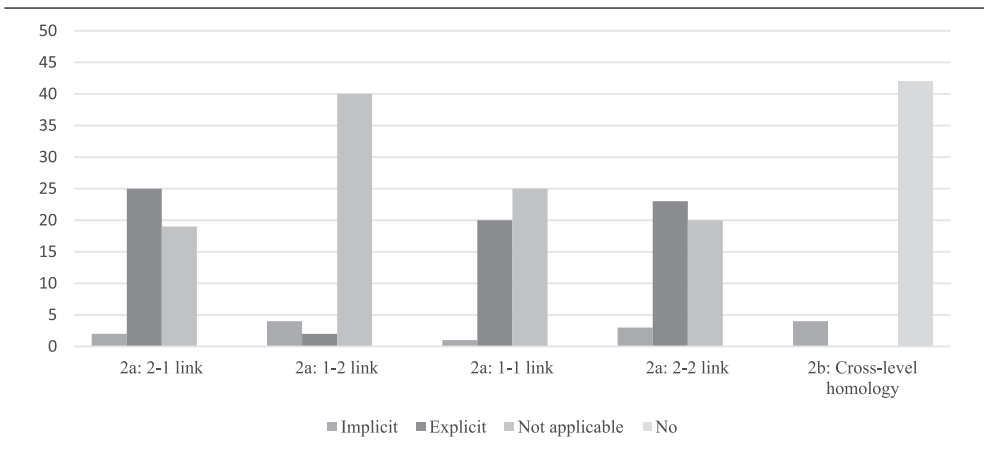
Figure 1
Conceptual Criteria

■ Yes ■ Implicit ■ Partial ▪ No ▪ Not applicable

Figure 2
Theoretical Criteria

■ Implicit ■ Explicit ▪ Not applicable ▪ No

The next set of criteria concerns the theoretical justification provided for the main hypothesized cross-level (2-1; 1-2) and single-level (1-1; 2-2) links in the model (Criterion 2a). Figure 2 shows that in a majority of the studies (25 out of 27) that hypothesized cross-level 2-1 links between HRM and outcomes, these links were explicitly justified by reference to core theories and/or previous empirical work, with a much smaller number of studies where these links were hypothesized but only implicitly justified (2 studies). In contrast, in terms of cross-level 1-2 links between employee attitudes, behaviors, and organizational outcomes, Figure 2 shows that these links were explicitly justified by reference to core theories and/or previous empirical work in only 2 out of 6 studies. In the vast majority of studies that included 1-1 and 2-2 links, on the other hand, these links were explicitly justified (20 out of 21 and 23

Figure 3
Measurement and Analysis Criteria
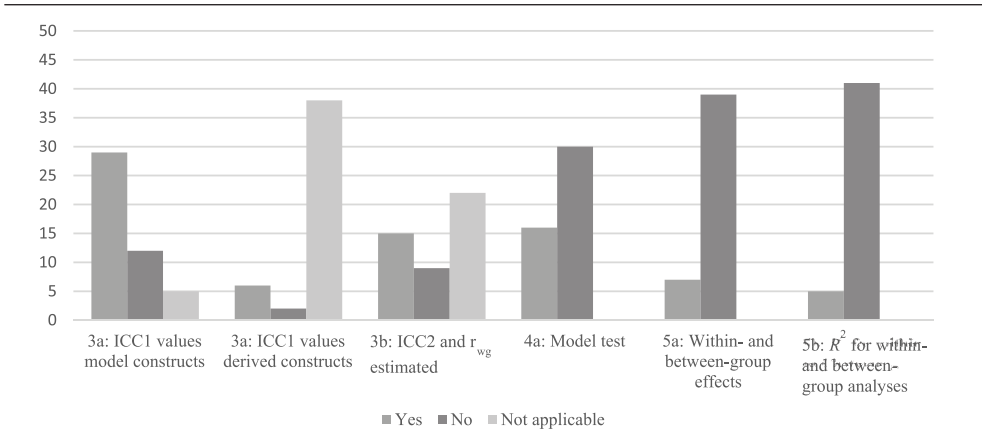
out of 26, respectively). Finally, a small number of studies (4) also mentioned cross-level homology, yet none of them either discussed this issue in detail or tested for it in their analyses (Criterion 2b).

Next, we assessed whether aggregation statistics were reported for all relevant study variables. Figure 3 indicates that ICC1 values for employee-rated model constructs were reported in 29 out of the 41 studies where it would have been appropriate to do so (Criterion 3a). Regarding studies that included derived between variables, ICC1 values for these variables were reported in 6 out of the 8 studies involved (Criterion 3a). In addition, Figure 3 shows that in a majority (15) of the 24 studies that included some form of employee-rated aggregate consensus-based Level 2 constructs, ICC2 and r_{wg} values were reported, while in 9 studies only some or none of these aggregation statistics were reported (Criterion 3b). Finally, in terms of the analysis criteria, the first of these criteria (4a) assesses whether the links hypothesized in extant studies were tested separately or simultaneously in the analysis. Figure 3 shows that in only a minority (16) of all extant studies were the hypothesized links tested simultaneously. Regarding the appropriate modeling and testing of within- and between-group effects (Criterion 5a), in only 7 studies were within and between effects tested separately by making use of MSEM to separate the between and within parts of all Level 1 variables. In addition, Figure 3 shows that only 5 of the studies reported separate R^2 values for within- and between-group analyses (Criterion 5b).

Discussion

Despite the growing emphasis placed on a multilevel perspective in HRM research, our review suggests that progress in this area has been relatively slow and uneven. In particular, our results suggest that the multilevel paradigm has not yet taken firm root in the field of HRM, with only a limited and uneven development of multilevel mediation models in the analysis of the HRM-outcomes relationship. This limited and uneven progress is reflected in the overall nature of the studies we examined as part of our review. Two points stand out in this respect.

First is the fact that nearly 40% (18) of the 46 studies we examined were proto-ML studies. While based on nested data and incorporating some elements of a multilevel perspective, such as the reporting of aggregation statistics, these studies ultimately remained rooted in a single-level approach focused on Level 2 relationships and based on the simple aggregation of Level 1 variables to the unit level of analysis. In essence, therefore, in these proto-ML studies, elements of the new multilevel paradigm simply tend to be grafted on to the historically dominant single-level approach to the analysis of the HRM-outcomes relationship. Hybrid studies of this kind that combine elements of two alternative research paradigms are indicative of a field in flux and are especially common in the early stages of the emergence of a new paradigm (Kuhn, 1962). In practice, this may not necessarily be a problem since, as noted, the results obtained using the simple aggregation procedures characteristic of a proto-ML approach can, under certain conditions, be similar to those that would be obtained using a more elaborate multilevel approach. However, researchers employing a simple aggregation approach would be well advised to check the robustness of their results by also analyzing their data using a full multilevel approach. As the new multilevel paradigm gains momentum and spreads, proto-ML studies can be expected to become less common and to taper off in favor of studies that are more centrally rooted in the new paradigm (Kuhn, 1962). For now, however, our analysis suggests that even though the multilevel paradigm has clearly gained ground in the field of HRM in recent years, in practice, there is still some way to go before this approach becomes fully dominant in HRM-outcomes research.

The second and more complex point to note concerns the extent to which the 28 SHRM and ECHRM non-proto-ML studies, or ML studies for short, covered in the review conform to the set of best practice criteria for the conduct of multilevel mediation research set out in our framework. As noted, although some of the ML studies we examined mentioned issues relating either to cross-level psychometric isomorphism or to homology, none of them considered these issues explicitly in their analyses (Criteria 1e and 2b). Apart from these two issues, however, many of the ML studies also fell short of best practice on a number of other criteria. As noted, only a minority of these studies, for example, tested cross-level mediation links in an integrated analytic framework (Criterion 4a), or explicitly distinguished within- from between-group effects, or reported relevant within- and between-group effect sizes separately (Criteria 5a and 5b). In contrast, a much larger proportion of these studies explicitly defined the Level 1 and 2 constructs included in their analyses (Criteria 1a and 1b), or provided more or less explicit justifications for the main hypothesized relationships in their models (Criterion 2a), or provided relevant aggregation statistics (Criteria 3a and 3b). In other words, our results suggest that in one way or another, virtually all extant multilevel studies of the HRM-outcomes relationship fall short of key best practice recommendations for the conduct of multilevel mediation research.

More generally, when evaluating extant ML studies against the set of best practice criteria, it is useful to distinguish between studies that adopted a partitioning approach and explicitly partitioned the variance in Level 1 constructs into within- and between-group components and studies that did not do so. Our analysis shows that only a relatively small minority of ML studies (7 out of 28) fell within the first category. By their very nature, these studies tend to be relatively sophisticated analytically and methodologically. In line with a partitioning approach, the majority of these studies not only explicitly model and test within and between effects separately using MSEM procedures but often also provide ICC1 values for all variables in the analysis and test hypothesized cross-level effects simultaneously in more

complex 2-1-2 and 2-1-1 models. In the vast majority of cases, however, this analytic and methodological sophistication is not necessarily matched by an equal level of conceptual and theoretical sophistication. In particular, despite adopting a partitioning perspective, virtually all the studies involved fail to provide a clear conceptualization of all the derived between Level 2 variables included in the analysis. As a result, the theoretical explanations that are provided for key hypothesized cross-level relationships in these studies often are not as convincing or compelling as they could be. In other words, the majority of studies involved here suffer from a lack of proper alignment between analytic/methodological and conceptual/ theoretical elements of the research.

The potential problems are different with respect to the second and much larger category of ML studies that did not adopt a partitioning approach. As noted, this second group includes 21 of the 28 non-proto-ML studies we examined. By their very nature, these studies tend to be less sophisticated analytically and methodologically in that they do not explicitly model and test within- and between-group effects separately. As a result, they fall short of most of the best practice measurement and analysis criteria highlighted in our framework. In compensation, however, many of the studies involved often provide fairly extensive definitions and conceptualizations of key Level 2 variables included in their multilevel models, as well as explicit theoretical explanations of core hypothesized cross-level relationships. In other words, many of the studies involved here tend to be stronger in conceptual/theoretical than in analytic/methodological terms.

Having said this, however, there are a number of additional points that should be noted about this second group of studies and about extant multilevel studies more generally. First, the conceptualization of key Level 2 constructs in many cases remains vague and unclear. In particular, as our results show, the actual nature of all the aggregate Level 2 model constructs included in these studies is often not clearly specified and/or is left implicit. Moreover, as we discuss more fully below, even though in the majority of studies the processes of emergence leading to within-group convergence in perceptions, attitudes, or behaviors of interest are specified, the explanations offered are often quite stylized, based on standard arguments linked, for example, to social information processing theory (Salancik & Pfeffer, 1978) or the ASA framework (Schneider, 1987). Equally important, the failure explicitly to partition within- and between-group variance in relevant Level 1 constructs, and the concomitant lack of information about effect sizes at different levels of analysis that is provided in the vast majority of the ML studies we examined, greatly limits substantive understanding of the HRM-outcomes relationship. In particular, it makes it difficult to gauge the extent to which HRM systems actually affect employee and organizational outcomes of interest at different levels of analysis and, therefore, to determine the true value and contribution of extant multilevel research in the area. This assessment is made all the more difficult by the potential biases introduced into the analysis by the failure explicitly to separate within- from between-group level effects in the vast majority of extant studies. And this, in turn, raises questions about the possible validity of many of the results and of the substantive conclusions about the HRM-outcomes relationship emerging from extant multilevel research in the area.

In sum, the results of our review show that proto-ML studies are still common in the analysis of the HRM-outcomes relationship. They also show that, in a variety of ways, the vast majority of multilevel studies proper in the area fall short of key current best practice

recommendations for the conduct of multilevel mediation research. Taken together, these two core findings point to the fact that, despite some clear advances, there still is a considerable way to go before the new paradigm takes firm root in the field of HRM. At the same time, however, our analysis also points to important ways in which the multilevel paradigm can contribute to the further development of both SHRM and ECHRM research. In particular, our review points to key underdeveloped aspects of the multilevel analysis of the HRM-outcomes relationship that, if systematically addressed, can make a substantial contribution to a deeper understanding of the effect of HR systems on both employee and organizational outcomes and of the mediating mechanisms involved. It is to these broader issues and directions for future research, therefore, that we now turn.

Looking Ahead

There are many ways in which a multilevel approach can contribute to a fuller understanding of the HRM-outcomes relationship and of the mediating mechanisms that underpin this relationship. Here we do not seek to provide a comprehensive overview of this potentially very large area of analysis. Nor is it our aim to provide an overall framework for the multilevel analysis of the HRM-outcomes relationship along the lines, for example, of the multilevel framework for SHRM research proposed by K. Jiang et al. (2013). Rather, building on our review of multilevel HRM-outcomes studies, our aim here is to highlight a number of important ways in which the further development and application of a multilevel approach can contribute to SHRM and ECHRM research. There are a number of specific issues that could usefully be considered here. These include, for example, issues about the conceptualization of aggregate Level 2 compositional constructs and about psychometric isomorphism and homology across levels of analysis or issues about the choice of nesting structure, such as the specification of salient supraindividual nesting units, and the incorporation of crosscutting nesting structures linked to the wider omnibus context in which individuals operate. All these issues have received only limited attention in extant studies and clearly deserve more systematic consideration in future research. For reasons of space, however, we necessarily have had to be selective in our coverage. Here, therefore, in order to make the discussion both more manageable and focused, and by way of illustration, we concentrate on a more limited number of core areas of analysis that, we believe, can prove particularly fruitful for the further substantive development of SHRM and ECHRM research.

As noted, to date, multilevel mediation studies of the HRM-outcomes relationship have focused exclusively on global and on additive or consensus-based compositional Level 2 constructs. That is to say, none of the multilevel studies we examined focused on more complex compilational forms of aggregation and explicitly incorporated compilation-based Level 2 independent, mediator, or dependent constructs in their models. In view of this, when considering ways in which the multilevel analysis of the HRM-outcomes relationship can be developed further, it is useful to distinguish between two broad ways forward. The first is to focus on developing and refining the analysis of standard multilevel models of the HRM-outcomes relationship based purely on global and compositional Level 2 constructs. And the second is to expand the scope of research in this area by extending the analysis to cover models that explicitly incorporate Level 2 independent, mediator, and/or outcome variables based on compilational forms of aggregation. Below we discuss these two lines of further inquiry.

Compositional models: Analysis of 2-1 links. A first key area for improvement in the multilevel mediation analysis of the HRM-outcomes relationship concerns the nature of the explanations that are provided for cross-level 2-1 links in SHRM and ECHRM research. From a substantive point of view, the central issue here is how to explain the potential effect of unit-level HR practices on the attitudes, experiences, and behaviors of employees within that unit (e.g., individual job satisfaction, stress, or citizenship behavior). As we have seen, the specific theories invoked in extant multilevel studies vary depending on the individual-level attitudes, experiences, or behaviors involved. Unfortunately, however, these theoretical explanations are seldom, if ever, integrated into a broader analysis that also takes into account the processes through which individual attitudes and behaviors coalesce and become manifest at the group level. The adoption of a partitioning perspective that explicitly distinguishes within- from between-group variance in Level 1 constructs is, we suggest, particularly important in this respect since such an approach can contribute to the development of a more integrated explanation of cross-level 2-1 effects in SHRM and ECHRM research.

In terms of theory development, one of the main implications of the partitioning perspective is that the explanation of the downward cross-level effect of a Level 2 variable, such as an HR system, on a Level 1 variable, such as employee job satisfaction, ultimately involves providing an adequate theoretical account of the observed relationship (covariation) between the HR system variable and group-level mean job satisfaction scores, representing the between-group component of the Level 1 job satisfaction variable. In turn, this theoretical account, we suggest, involves explaining two main processes. First, it involves explaining how and why employees' exposure to a given set of HR practices affects their individual job satisfaction. And second, it requires explaining how, over time, individual job satisfaction coalesces to the group level.

In order to gain a better understanding of the cross-level effect of HR practices on individual attitudes and behaviors of interest in SHRM and ECHRM research, therefore, the two-step explanatory approach outlined above suggests it is important to pay more systematic attention to two key issues. The first issue concerns the extent to which unit-level HR systems actually have a more or less uniform effect on individual employee attitudes and behaviors. A good example of this kind of analysis is provided by the study by Den Hartog et al. (2013) included in our review. In this study, the cross-level relationship between management-rated high performance work systems (HPWS) at Level 2 and employee-rated HPWS at Level 1 was hypothesized and shown to be moderated by individuals' perceptions of effective supervisor communication at Level 1.

This kind of moderated mediation analysis (Bullock, Green, & Ha, 2010; Hayes, 2015; Preacher, Rucker, & Hayes, 2007), however, is more the exception than the rule in the multilevel studies we examined (Shin et al., in press; Zhong, Wayne, & Liden, 2016). Based, for example, on social exchange theory or on labor process arguments, the vast majority of these studies assume that all employees experience and/or react to the HR practices they are exposed to in a similar positive or negative way and that these practices, therefore, tend to have a homogeneous either positive or negative effect on their attitudes and behaviors (Liao et al., 2009; Messersmith et al., 2011; Van De Voorde, Veld, & Van Veldhoven, 2016; Wood, Van Veldhoven, Croon, & de Menezes, 2012). There are, however, a number of variables at both the individual and the unit level of analysis that could usefully be examined as potential moderators of the HR practices–employee attitudes and behavior relationship. At the

individual level, these include, for example, the work expectations and orientations of employees, the extent to which they trust management, their attributions of why management has introduced the practices in the first place, and the evenness with which practices are actually implemented by first-line supervisors within groups (Boxall & Purcell, 2008; Nishii et al., 2008). In terms of trust, for example, employees who perceive management as less trustworthy can be expected to view many HR practices, such as security of employment guarantees, with more suspicion than employees who view management as more trustworthy and genuinely concerned with their well-being. As a result, low trust employees are likely to respond less positively than high trust employees to the use, for example, of high performance work practices. Hence, in line with social exchange theory, low trust employees are less likely to reciprocate the use of HPWS with higher levels of loyalty and commitment to the organization or by engaging in more extensive forms of citizenship behavior.

At the group level, potential moderators of the HR practices–employee attitudes and behavior relationship include, for example, type of leadership, team cohesion and team task complexity, or HRM process features (S. Chang et al., 2014; K. Jiang, Chuang, & Chiao, 2015; Katou, Budhwar, & Patel, 2014). At the level of the work group, for instance, the potential impact of HR practices on group members is likely to vary depending on how effectively those practices are actually implemented by first-line supervisors and, therefore, on the experience and leadership qualities of supervisors and on their ability and motivation to translate and implement HR practices at the local level. More generally, all the potential moderators identified above, at both the individual and the group level, are linked to major ongoing debates in the HRM literature about the conditions that are likely to affect the extent to which HR practices will have a positive or negative effect on outcomes of interest in SHRM and ECHRM research (Jackson et al., 2014; Jensen & Van De Voorde, 2016; Peccei et al., 2013). The systematic inclusion of these type of moderators can, therefore, greatly contribute to the further development of research in the area by providing a fuller understanding of the initial 2-1 cross-level links in multilevel models and the conditions under which unit-level HR practices are more or less likely to affect employee attitudes and behaviors to begin with.

The two-step explanatory approach outlined above suggests that to improve understanding of the cross-level effect of HR practices on individual attitudes and behaviors, it is also important to pay more systematic attention to a second key issue. This concerns the processes of emergence that foster the development of consensus and convergence in such attitudes and behaviors within groups over time. As noted, all the multilevel studies we examined simply tend to assume that some form of convergence in attitudes and behaviors actually takes place within groups. This presumed convergence is commonly explained by reference to generic theories capable of accounting for the possible development of greater within-group consensus in attitudes and behaviors of interest, such as social information processing theory (Salancik & Pfeffer, 1978) or ASA arguments (Schneider, 1987). None of the studies, however, actually examine the specific dynamics of emergence. This includes, for example, the way in which processes of convergence actually unfold over time within groups and the speed and outcome of these processes.

Clearly, many factors may come into play here, such as the frequency of interaction between group members, the degree of demographic diversity within the group, and the nature and composition of internal social networks (Harrison & Klein, 2007; Kozlowski, Chao, Grand, Braum, & Kuljanin, 2013; Kozlowski & Klein, 2000). Social influence

processes and, therefore, attitude convergence can, for example, be expected to take place faster and more easily amongst individuals who work in close-knit teams and interact frequently with each other than amongst employees operating in more loosely structured work groups with fewer opportunities for face-to-face interaction. Also important here, however, is the internal heterogeneity of the group in terms of key variables, such as trust in management, that may moderate the effect of HR practices on individual attitudes and behaviors of interest (e.g., organizational commitment). The greater the heterogeneity of the group in this respect, the more likely it is that the introduction of new HR practices will have a nonuniform effect on employee commitment. Other things equal, therefore, the greater the heterogeneity, the longer it is likely to take and/or the more difficult it is likely to be for a high degree of consensus on commitment eventually to emerge within the group. The overall level of commitment of the group might also be expected to be lower than if, as a result of consistently high levels of trust in management amongst group members, the HR practices had a uniformly strong effect on employee commitment to begin with.

Finally, in this context, it is also important to consider the precise nature of the aggregate Level 2 consensus constructs involved. Particularly important in this respect is whether the constructs in question are conceptualized in terms of direct consensus or referent-shift forms of composition. This is the difference, for example, between a simple aggregate organizational commitment construct that indexes and then averages to the group level individual members' own reported level of commitment and a collective commitment construct that indexes and then averages to the group level individuals' perceptions of the group's, rather than of their own, level of commitment to the organization (Chan, 1998; Wallace et al., 2016). This distinction is important because the factors affecting both the emergence and then the subsequent process of convergence of collective Level 2 constructs may well be different from those affecting simple aggregate constructs. For example, the simple average level of organizational commitment in a group may depend on the extent to which supervisors consistently share information with or provide developmental opportunities for individuals in the group. Collective group commitment, on the other hand, may also depend on the extent to which supervisors emphasize team-based learning activities or share information through collective team briefings, thereby enabling group members more easily to interact and share their views about the organization. In the absence of such group practices, it may be more difficult for collective organizational commitment to emerge within a group. In addition, it is worth noting that consistent with research on, for example, organizational climate, collective and simple aggregate Level 2 constructs may also have a different effect on both employee and organizational outcomes (Wallace at al., 2016). Hence, the conceptualization of Level 2 aggregate constructs is central to the analysis of multilevel relationships and, therefore, deserves particular attention in future multilevel HRM-outcomes research.

More generally, therefore, SHRM and ECHRM research can benefit from a more fine-grained and dynamic analysis of the processes of emergence (Kozlowski & Klein, 2000) through which, over time, individual employee attitudes and behaviors coalesce and ultimately manifest at group level. When combined with a more systematic consideration of the way exposure to HR practices affects individual employee attitudes and behaviors to begin with, of the potential moderators of this relationship, and of the meaning and nature of Level 2 aggregate constructs, such a more dynamic analysis can, we suggest, provide a better and more integrated understanding of the how, why, and when of the downward cross-level

impact of HRM on employee outcomes that are of central importance to both SHRM and ECHRM research.

Compositional models: Analysis of 1-2 links. A second key area for improvement in the multilevel analysis of the HRM-outcomes relationship concerns the analysis of upward cross-level (1-2) links in full-mediation 2-1-2 models in SHRM research. In substantive terms, 1-2 links are important because they are at the core of key ongoing debates in the HRM-outcomes literature about the extent to which employee attitudes, experiences, and behaviors actually mediate the effect of unit-level HR practices on unit-level performance outcomes of interest. Drawing on the AMO framework, for example, these include debates about the extent to which particular systems of HR practices (e.g., HPWS) affect the AMO of employees to perform on their job (i.e., the 2-1 links discussed above) and the extent to which this, in turn, affects key aspects of organizational performance (i.e., the 1-2 links of interest here; Jackson et al., 2014; K. Jiang et al., 2012). Importantly, it also includes ongoing debates about the extent to which HR practices affect different aspects of employee well-being (e.g., happiness vs. health-related well-being, such as job satisfaction vs. stress) and the extent to which these different aspects of well-being, in turn, affect different elements of organizational performance (i.e., the 1-2 effects of interest here; Jackson et al., 2014; Van De Voorde et al., 2012). In turn, the latter debates are part of a wider ongoing discussion in the ECHRM and SHRM literature about a mutual gains versus competing outcomes interpretation of the impact of HR systems on key aspects of both employee well-being and organizational performance (Peccei et al., 2013).

The analysis of these core bottom-up 1-2 links bears important similarities to the analysis of downward 2-1 links discussed above. Here too a type of two-step explanatory approach applies. More specifically, from a substantive point of view, the central issue with respect to 1-2 links is how to explain the effect of individual employee attitudes, experiences, and behaviors on aggregate unit-level outcomes of interest. The issue, for example, is how to explain the effect of employee job satisfaction (a Level 1 well-being variable) on unit-level productivity or customer satisfaction (a Level 2 outcome). Once again, in the case of compositional models, this theoretical account, we suggest, involves explaining two main processes. First, it involves explaining how, over time, individual job satisfaction coalesces to the group level. And second, it involves explaining how the aggregate-level component of job satisfaction (i.e., the derived between group score of the individual-level variable) is related to the outcome of interest at the unit level.

The key problems and issues involved in the analysis of processes of emergence here are similar to those discussed above in relation to the analysis of the upward component of 2-1 links. Here, therefore, it is sufficient to focus on the second element involved in the explication of 1-2 links in 2-1-2 models. This involves accounting for the relationship between the aggregate component of hypothesized Level 1 attitudinal and behavioral variables and the unit-level outcomes of interest. In particular, this second step involves explaining how aggregate (derived between group) job satisfaction, for example, actually affects unit-level productivity or customer satisfaction. Importantly, because the relationships of interest here involve purely Level 2 variables, explanations of these relationships might most usefully focus on processes and mechanisms that also operate at the group or unit level. For example, drawing on social exchange and citizen behavior theory (Whitman, Van Rooy, & Viswesvaran,

2010), high aggregate levels of job satisfaction within a work group can be expected to contribute to the development of more positive and harmonious group relations which, in turn, is likely to contribute to group performance by, for instance, enhancing the ability of the group as a whole to provide more effective service to customers. We return to the issue of Level 2 theories and mechanisms below. Here it is sufficient to note that the specific Level 2 theories and arguments that might be deployed will necessarily vary depending on the particular Level 2 relationships and outcomes involved. More generally, however, SHRM research would benefit from a more detailed dynamic analysis of 1-2 relationships in 2-1-2 mediation models. In particular, it would benefit from paying more systematic attention, first, to the processes of emergence whereby individual attitudes, experiences, and behaviors eventually aggregate and coalesce at the group level and, second, to the Level 2 processes and mechanisms whereby collective employee attitudes and behaviors ultimately influence important Level 2 outcomes of interest to SHRM researchers.

Compositional models: Combining explanations at different levels of analysis. As we have seen, central to extant 2-1-2 mediation studies of the HRM-outcomes relationship is the idea that individual employee attitudes, experiences, and behaviors are the core mechanism through which unit-level HR practices affect unit-level outcomes. As noted by K. Jiang et al. (2013), however, in addition to these micromechanisms operating at the individual level, there are other higher-order mechanisms operating directly at the organizational level that may mediate the unit-level HRM-outcomes relationship (see also Jackson et al., 2014). Particularly important in this respect, for example, are various types of organizational capabilities, such as adaptive capability (L. Q. Wei & Lau, 2010), organizational ambidexterity (Patel, Messersmith, & Lepak, 2013), and knowledge integration (Collins & Smith, 2006). Drawing on the resource-based view, these organizational-level dynamic capabilities have been hypothesized and shown to play a significant mediating role in various aspects of the HRM-outcomes relationship.

Despite the potential importance of these macrolevel mediating mechanisms operating directly at the unit level of analysis, none of the 2-1-2 mediation studies we examined included any systematic discussion of possible supraindividual mechanisms through which Level 2 HRM systems might directly affect higher-order outcomes also at Level 2. Even studies that included a direct macro-macro pathway between HRM and outcome variables at Level 2, in what are essentially Coleman-type multilevel models (Coleman, 1990), failed to explicate the nature of the 2-2 pathway involved (e.g., Wood et al., 2012). More generally, therefore, we suggest that given the potential importance of macrolevel mechanisms in multilevel analyses of the HRM-outcomes relationship, greater attention needs to be paid to such mechanisms when testing 2-1-2 models in future SHRM research. In particular, the aim is to incorporate macrolevel mechanisms into existing 2-1-2 models so as to be able to examine the effects of mediating mechanisms operating at the individual and at the unit level simultaneously. Such an exercise can, we believe, significantly contribute to SHRM research by enabling researchers systematically to assess the relative importance of mediating mechanisms and, therefore, of explanations of the HRM-outcomes relationship that operate at different levels of analysis.

Compilational models and the analysis of configurations. A final and possibly more fundamental way of enriching SHRM and ECHRM research is by explicitly focusing on more

complex compilational constructs as part of the multilevel analysis of the HRM-outcomes relationship. Several lines of inquiry might prove useful in this respect. One possibility, for example, is to consider not only the mean score of aggregate Level 2 variables in multilevel models but also the pattern of dispersion of these variables (e.g., the consistency of individuals' responses within groups). In general, for example, mean unit-level HR perception scores may be positively related to employee job satisfaction, but the strength of the relationship may depend on the consistency in HRM perceptions across individuals within groups (i.e., HRM strength; Bowen & Ostroff, 2004). In addition, diversity in well-being among unit members may affect group processes and performance indicators (Barsade, Ward, Turner, & Sonnenfield, 2000). Here we focus on what we believe to be a particularly promising line of inquiry involving configurational forms of compilation. Specifically, we focus on a set of issues linked to the use of configurational variables in the analysis of the HRM-outcomes relationship that can significantly contribute to the further development and integration of SHRM and ECHRM research. By way of illustration, we briefly consider three main ways in which this can be done, namely, by incorporating Level 2 compilation-based configurational constructs into 2-1-2 models as independent, mediator, and outcome variables, respectively.

Before addressing these issues, it is worth noting that, in the context of multilevel research, it is important to distinguish between two broad types of Level 2 configurations, global and compilational, respectively. Global Level 2 configurations are based on the overall summary scores of a single respondent, such as a manager's ratings of the unit's HR practices. Compilational Level 2 configurations, on the other hand, are based on the aggregated ratings provided by multiple Level 1 raters, such as the HR practices perceptions of employees within the unit. Although there are different ways of constructing the latter type of Level 2 configurations from individual Level 1 scores, these configurations are all based on a compilational logic in that they ultimately index variables at the supraindividual level that cannot be decomposed or have a direct corresponding analogue at the individual level. More generally, therefore, all (nonglobal) independent, mediator, or dependent configurational variables in 2-1-2 models that are constructed through some form of aggregation of Level 1 variables to the unit level are necessarily compilational in nature.

A key debate in the HRM literature concerns whether to conceptualize HR systems in either simple additive or more complex configurational terms (Chadwick 2010; Delery, 1998; Gerhart, 2013). Unlike additive conceptualization, configurational conceptualizations focus on potential synergies or interactions between HR practices and on the effects, therefore, that different combinations of practices have on outcomes. In other words, the emphasis of configurational conceptualizations is on different types of HR systems reflecting different patterns, combinations, or profiles of HR practices. Although configurational approaches to the conceptualization of HR systems have attracted increasing attention in the HRM literature (Delery & Doty, 1996; Hauff, Alewell, & Hansen, 2014; Toh, Morgeson, & Campion, 2008), additive conceptualizations still dominate multilevel SHRM and ECHRM research.

Conceptualizations of HR systems in configurational terms and their inclusion as Level 2 independent variables in multilevel models can, however, contribute to a number of important areas of analysis that are central to major ongoing debates in SHRM and ECHRM. One such area, for example, concerns the analysis of HRM segmentation (Jackson et al., 2014; K. Jiang et al., 2013; Lepak & Snell, 2002). Particularly important in this respect is the possibility, through the inclusion of Level 2 HR system configurations in multilevel models, of

exploring more systematically whether organizations do indeed adopt different combinations of HR practices for different employee groups depending, for example, not only on the strategic importance and contribution of the group to the organization but also on the occupational composition of the group and its position in the functional and hierarchical division of labor of the organization (Buller & McEvoy, 2012; Huselid & Becker, 2011; Lepak & Snell, 2002; Ostroff & Bowen, 2000). Equally important is the possibility of exploring more systematically the conditions under which organizations are more or less likely to adopt internally differentiated HR architectures, as well as the consequences of HRM segmentation for both employees and organizations. In terms of the consequences of segmentation, a particularly important issue needing further systematic attention concerns the extent to which different systems of HR practices are actually required to maximize the job performance and well-being of different groups of employees and the conditions under which this is more or less likely to be the case.

This type of analysis can be refined further by developing various kinds of profiles or configurations of employee outcomes of interest and then introducing these into multilevel models at Level 2 as employee-based compilational configurations. Of particular relevance here is the major ongoing debate in the ECHRM literature concerning the extent to which HR systems (e.g., HPWS or high involvement HR systems) actually have a positive or negative effect on different aspects of employee well-being, such as happiness- versus health-related well-being (Peccei et al., 2013; Van De Voorde et al., 2012). These different dimensions of well-being can usefully be used to construct distinct employee well-being profiles that, when aggregated to the unit level, can then be related to different unit-level HR system configurations. This would enable researchers to contribute to the analysis of ECHRM by gaining a better understanding of the particular types of unit-level HR systems that are most likely simultaneously to enhance the happiness- and health-related aspects of well-being and that, therefore, might be most desirable from an employee perspective. Equally important, this type of analysis would also allow for a more systematic consideration of the particular combination of HR practices that are likely to have a negative effect on both aspects of employee well-being. In turn, these configurations can be compared with those that may have a mixed effect by, for example, enhancing one aspect of well-being (e.g., job satisfaction linked to high pay and autonomy) but at the expense of other dimensions (e.g., higher stress linked to high job demands).

More generally, therefore, the inclusion of compilation-based well-being profiles into multilevel models of the HRM-outcomes relationship can make a substantial contribution to the analysis of well-being trade-offs linked to the adoption of different particular types of HR system configurations by organizations. Central to such a novel line of inquiry, however, is also the need to develop a theoretical understanding of how and why particular configurations of HR practices may be associated with given configurations of employee well-being, as well as of the conditions under which particular configurational relationships hold.

The multilevel configurational analyses outlined above can also usefully be extended to the simultaneous consideration of employee and organizational outcomes of interest. In particular, the aim here would be to contribute to the analysis of the HRM-outcomes relationship by combining some of the core concerns of SHRM and ECHRM research and, in the words of Jackson et al., "Take seriously the concerns of multiple stakeholders" (2014: 31). One fruitful way of doing this, we suggest, is by first aggregating employee well-being data to the

unit level. These (mean) unit-level employee well-being scores can then be combined with various aspects of unit performance (e.g., global productivity scores) to yield distinct aggregate, unit-level well-being performance profiles that can then be systematically related to different HR system configurations. An important contribution of this type of unit-level compilational/configurational analysis is that it would enable researchers systematically to explore which specific types of unit-level HR system configurations are more likely to contribute simultaneously to enhance both employee well-being and organizational performance and, therefore, are more likely to be associated with win-win situations that directly benefit both employees and management (Peccei et al., 2013; Van De Voorde et al., 2012). Importantly, this type of more complex configurational analysis can also contribute to the analysis of the HRM-outcomes relationship by enabling researchers to identify types of HR systems that produce either mixed or systematically poorer outcomes for management and employees, reflecting different possible combinations of win-lose situations for the parties involved. As such, the proposed analysis can also shed some light on an important area of growing debate in the HRM literature, namely, the question of the trade-offs that may exist between organizational performance and employee well-being and the role that HR systems play in such trade-offs (Peccei, 2004; Van De Voorde et al., 2012).

In sum, the type of multilevel compilation-based configurational analyses highlighted here represent an important departure from the dominant compositional approaches used in extant multilevel studies of the HRM-outcomes relationship. Superficially, multilevel configurational models resemble the proto-ML models discussed above in that all the key variables in configurational analyses, like those in proto-ML models, explicitly reside at the supraindividual level. The core difference, however, is that the key Level 2 variables in configurational models, unlike those in proto-ML models, are compilational rather than compositional in nature. As a result, partitioning issues, which are important but tend to be ignored in proto-ML models, do not necessarily apply to the analysis of multilevel relationships in configurational models. More generally, the configurational analyses outlined above can, we suggest, substantively advance HRM-outcomes research by contributing to a better understanding of the effect that potentially complex systems of HR practices can simultaneously have on both employee outcomes and organizational performance. As such, the proposed lines of analyses can contribute to a greater integration of SHRM and ECHRM thinking and research.

Conclusions

In this article, we have provided the first systematic review and evaluation of the development and application of the multilevel paradigm in HRM research focusing, in particular, on multilevel mediation studies of the HRM-outcomes relationship. Our results showed that even though a growing number of SHRM and ECHRM studies of the HRM-outcomes relationship have indeed adopted a multilevel approach, progress in the area has been limited and uneven. First, proto-ML studies that in key respects remain rooted in a single-level paradigm continue to be common. Second, as part of our analysis, we evaluated extant studies against a number of important best practice theoretical and methodological criteria for the conduct of multilevel research derived from the wider literature. Although developed specifically with multilevel HRM-outcomes research in mind, the set of best practice criteria can, we hope, also serve as a useful guide for multilevel management researchers more generally.

When used to evaluate extant multilevel HRM-outcomes studies, the analysis revealed that all the studies involved fell short of key best practice recommendations on a number of dimensions. On the basis of this evaluation, we then identified a number of important areas for improvement in the application of a multilevel approach to the analysis of the HRM-outcomes relationship. If systematically addressed, these areas for improvement can, we believe, significantly contribute to the further theoretical and methodological development of SHRM and ECHRM research.

Beyond these specific areas of best practice improvement, we also highlighted a number of more fundamental ways in which the further refinement and extension of multilevel mediation models of the HRM-outcomes relationship can make a significant substantive contribution to the further development of SHRM and ECHRM research. We illustrated this by reference to specific substantive issues related to the analysis of cross-level relationships and the operation of mediating mechanisms at different levels of analysis in compositional models. Importantly, we also extended the analysis to a consideration of compilational models. In particular, we highlighted the contribution that a multilevel configurational approach can make to SHRM and ECHRM research by focusing attention on an analysis of the effect that different types of unit-level HR systems may have on complex combinations of both employee and organizational outcomes and of the mediating mechanisms involved.

Beyond the issues identified above, there clearly are many other important areas of analysis linked to the application of the multilevel paradigm that need to be explored by scholars interested in contributing to SHRM and ECHRM research. At the same time, it is also important to recognize that addressing some of the issues and areas for improvement identified here is likely to involve very significant theoretical, methodological, and practical challenges. We hope, however, that the present study can serve as an initial point of departure and reflection for future research in the area and for the application of the multilevel paradigm in the field of HRM more generally.

References

References marked with an asterisk indicate multilevel studies included in the analysis.

Aguinis, H., Boyd, B., Pierce, C., & Short, J. 2011. Walking new avenues in management research methods and theories: Bridging micro and macro domains. *Journal of Management*, 37: 395-403.

*Allen, M., Ericksen, J., & Collins, C. J. 2013. HRM, employee exchange relationships, and performance in small businesses. *Human Resource Management*, 52: 153-174.

Arthur, J. B. 1994. Effects of human resource systems on manufacturing performance and turnover. *Academy of Management Journal*, 37: 670-687.

*Aryee, S., Walumbwa, F., Seidu, E., & Otaye, L. 2012. Impact of HPWS on individual- and branch-level performance: Test of a multilevel model of intermediate linkages. *Journal of Applied Psychology*, 97: 287-300.

*Aryee, S., Walumbwa, F., Seidu, E., & Otaye, L. 2016. Developing and leveraging human capital resource to promote service quality: Testing a theory of performance. *Journal of Management*, 42: 480-499.

Bal, P. M., Kooij, D., & De Jong, S. B. 2013. How do developmental and accommodative HRM enhance employee engagement and commitment? The role of psychological contract and SOC strategies. *Journal of Management Studies*, 50: 545-572.

*Baluch, A. M., Salge, T. O., & Piening, E. P. 2013. Untangling the relationship between HRM and hospital performance: The mediating role of attitudinal and behavioural HR outcomes. *The International Journal of Human Resource Management*, 24: 3038-3061.

*Barrick, M. R., Thurgood, G. R., Smith, T. A., & Courtright, S. H. 2015. Collective organizational engagement: Linking motivational antecedents, strategic implementation, and firm performance. *Academy of Management Journal*, 58: 111-135.

Barsade, S., Ward, A., Turner, J., & Sonnenfeld, J. 2000. To your heart's content: A model of affective diversity in top management teams. *Administrative Science Quarterly*, 45: 802-836.

Becker, B. E., & Huselid, M. A. 1998. High performance work systems and firm performance: A synthesis of research and managerial implications. In K. M. Rowlands & G. R. Ferris (Eds.), *Research in personnel and human resource management*, vol. 16: 53-101. Greenwich, CT: JAI Press.

Bliese, P. D. 2000. Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions*: 349-381. San Francisco: Jossey-Bass.

Boselie, P., Dietz, G., & Boon, C. 2005. Commonalities and contradictions in HRM and performance research. *Human Resource Management Journal*, 15(3): 67-94.

Boxall, P., & Purcell, P. 2008. *Strategy and HRM*. Basingstoke, England: Palgrave Macmillan.

Bowen, D. E., & Ostroff, C. 2004. Understanding HRM–firm performance linkages: The role of the "strength" of the HRM system. *Academy of Management Review*, 29: 203-221.

Buller, P. F., & McEvoy, G. M. 2012. Strategy, human resource management and performance: Sharpening line of sight. *Human Resource Management Review*, 22: 43-56.

Bullock, J. G., Green, D. P., & Ha, S. E. 2010. Yes, but what's the mechanism? (don't expect an easy answer). *Journal of Personality and Social Psychology*, 98: 550-558.

Burke, M. J., Finkelstein, L. M., & Dusig, M. S. 1999. On average deviation indices for estimating interrater agreement. *Organizational Research Methods*, 2: 49-68.

Chadwick, C. 2010. Theoretic insights on the nature of performance synergies in human resource systems: Toward greater precision. *Human Resource Management Review*, 20: 85-101.

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

*Chang, P. C., & Chen, S. J. 2011. Crossing the level of employee's performance: HPWS, affective commitment, human capital, and employee job performance in professional service organizations. *The International Journal of Human Resource Management*, 22: 883-901.

Chang, S., Jia, L., Takeuchi, R., & Cai, Y. 2014. Do high-commitment work systems affect creativity? A multilevel combinational approach to employee creativity. *Journal of Applied Psychology*, 99: 665-680.

Chen, G., Bliese, P. D., & Mathieu, J. E. 2005. Conceptual framework and statistical procedures for delineating and testing multilevel theories of homology. *Organizational Research Methods*, 8: 375-409.

Chen, G., Mathieu, J., & Bliese, P. 2005. A framework for conducting multi-level construct validation. In F. Yammarino & F. Dansereau (Eds.), *Multi-level issues in organizational behavior and processes*, vol. 3: 273-303. Bingley, England: Emerald Group.

*Chuang, C. H., & Liao, H. 2010. SHRM in service context: Taking care of business by taking care of employees and customers. *Personnel Psychology*, 63: 153-196.

Coleman, J. 1990. *Foundations of social theory*. Cambridge, MA: Harvard University Press.

*Collins, C., & Smith, K. 2006. Knowledge exchange and combination: The role of HR practices in the performance of high-technology firms. *Academy of Management Journal*, 49: 544-560.

Combs, J., Liu, Y., Hall, A., & Ketchen, D. 2006. How much do high-performance work practices matter? A meta-analysis of their effects on organizational performance. *Personnel Psychology*, 59: 501-528.

Croon, M. A., van Veldhoven, M., Peccei, R., & Wood, S. J. 2014. Researching individual well-being and performance in context: Multilevel mediational analysis for bathtub models. In M. van Veldhoven & R. Peccei (Eds.), *Well-being and performance at work: The role of context*: 129-154. Hove, England: Psychology Press.

Dansereau, F., & Yammarino, F. J. 2000. Within and between analysis: The variant paradigm as an underlying approach to theory building and testing. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions*: 425-466. San Francisco: Jossey-Bass.

Delery, J. E. 1998. Issues of fit in strategic human resource management: Implications for research. *Human Resource Management Review*, 8: 289-309.

Delery, J. E., & Doty, D. H. 1996. Modes of theorizing in SHRM: Tests of universalistic, contingency, and configurational performance predictions. *Academy of Management Journal*, 39: 802-835.

*Dello Russo, S., Mascia, D., & Morandi, F. in press. Individual perceptions of HR practices, HRM strength and appropriateness of care: A meso, multilevel approach. *The International Journal of Human Resource Management*. doi:10.1080/09585192.2016.1165276

*Den Hartog, D., Boon, C., Verburg, R., & Croon, M. 2013. HRM, communication, satisfaction, and perceived performance: A cross-level test. *Journal of Management*, 39: 1637-1665.

Dyer, N. G., Hanges, P. J., & Hall, R. J. 2005. Applying multilevel confirmatory factor analysis techniques to the study of leadership. *The Leadership Quarterly*, 16: 149-167.

*Elorza, U., Aritzeta, A., & Ayestarán, S. 2011. Exploring the black box in Spanish firms: The effect of the actual and perceived system on employees' commitment and organizational performance. *The International Journal of Human Resource Management*, 22: 1401-1422.

*Elorza, U., Harris, C., Aritzeta, A., & Balluerka, N. 2016. The effect of management and employee perspectives of high-performance work systems on employees' discretionary behaviour. *Personnel Review*, 45: 121-141.

*Flinchbaugh, C., Li, P., Luth, M. T., & Chadwick, C. 2016. Team-level high involvement work practices: Investigating the role of knowledge sharing and perspective taking. *Human Resource Management Journal*, 26(2): 134-150.

Fulmer, C. A., & Ostroff, C. 2016. Convergence and emergence in organizations: An integrative framework and review. *Journal of Organizational Behavior*, 37: S122-S145.

*Gardner, T. M., Wright, P. M., & Moynihan, L. M. 2011. The impact of motivation, empowerment, and skill-enhancing practices on aggregate voluntary turnover: The mediating effect of collective affective commitment. *Personnel Psychology*, 64: 315-350.

*Gelade, G. A., & Ivery, M. 2003. The impact of human resource management and work climate on organizational performance. *Personnel Psychology*, 56: 383-404.

Gerhart, B. 2013. Research on human resources and effectiveness: Some methodological challenges. In J. Paauwe, D. Guest, & P. M. Wright (Eds.), *HRM and performance: Achievements and challenges*: 149-172. Oxford, UK: Wiley-Blackwell.

Godard, J. 2001. Beyond the high-performance paradigm? An analysis of variation in Canadian managerial perceptions of reform programme effectiveness. *British Journal of Industrial Relations*, 39: 25-52.

*Gong, Y., Law, K. S., Chang, S., & Xin, K. R. 2009. HRM and firm performance: The differential role of managerial affective and continuance commitment. *Journal of Applied Psychology*, 94: 263-275.

Guest, D. 1999. HRM—The workers' verdict. *Human Resource Management Journal*, 9(3): 5-25.

Guest, D. 2011. Human resource management and performance: Still searching for some answers. *Human Resource Management Journal*, 21(1): 3-13.

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.

Hauff, S., Alewell, D., & Hansen, N. K. 2014. HRM systems between control and commitment: Occurrence, characteristics and effects on HRM outcomes and firm performance. *Human Resource Management Journal*, 24(4): 424-441.

Hayes, A. F. 2015. An index and test of linear moderated mediation. *Multivariate Behavioral Research*, 50: 1-22.

Hitt, M. A., Beamish, P. W., Jackson, S. E., & Mathieu, J. E. 2007. Building theoretical and empirical bridges across levels: Multilevel research in management. *Academy of Management Journal*, 50: 1385-1399.

Hofmann, D. A. 1997. An overview of the logic and rationale of hierarchical linear models. *Journal of Management*, 23: 723-744.

Hofmann, D. A., Griffin, M. A., & Gavin, M. B. 2000. The application of hierarchical linear modeling to organizational research. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions*: 467-511. San Francisco: Jossey-Bass.

*Hong, Y., Liao, H., Raub, S., & Han, J. H. 2016. What it takes to get proactive: An integrative multilevel model of the antecedents of personal initiative. *Journal of Applied Psychology*, 101: 687-701.

*Huang, L. C., Ahlstrom, D., Lee, A. Y. P., Chen, S. Y., & Hsieh, M. J. 2016. HPWS, employee well-being, and job involvement: An empirical study. *Personnel Review*, 45: 296-314.

Huselid, M. A. 1995. The impact of HRM practices on turnover, productivity, and corporate financial performance. *Academy of Management Journal*, 38: 635-672.

Huselid, M. A., & Becker, B. E. 2011. Bridging micro and macro domains: Workforce differentiation and SHRM. *Journal of Management*, 37: 421-428.

Jackson, S. E., Schuler, R. S., & Jiang, K. 2014. An aspirational framework for strategic human resource management. *The Academy of Management Annals*, 8: 1-56.

James, L. R., Demaree, R. G., & Wolf, G. 1984. Estimating within-group interrater reliability with and without response bias. *Journal of Applied Psychology*, 69: 85-98.

Jensen, J. M., Patel, P. C., & Messersmith, J. G. 2013. HPWS and job control: Consequences for anxiety, role over-load, and turnover intentions. *Journal of Management*, 39: 1699-1724.

Jensen, J. M., & Van De Voorde, K. 2016. Reconciling the dark side of HPWS. In N. M. Ashkanasy, R. J. Bennett, & M. J. Martinko (Eds.), *Understanding the high performance workplace: The line between motivation and abuse*: 85-102. New York: Routledge.

*Jiang, J., Wang, S., & Zhao, S. 2012. Does HRM facilitate employee creativity and organizational innovation? A study of Chinese firms. *The International Journal of Human Resource Management*, 23: 4025-4047.

*Jiang, K., Chuang, C. H., & Chiao, Y. C. 2015. Developing collective customer knowledge and service climate: The interaction between service-oriented high-performance work systems and service leadership. *Journal of Applied Psychology*, 100: 1089-1106.

Jiang, K., Lepak, D. P., Hu, J., & Baer, J. C. 2012. How does HRM influence organizational outcomes? A meta-analytic investigation of mediating mechanisms. *Academy of Management Journal*, 55: 1264-1294.

Jiang, K., Takeuchi, R., & Lepak, D. 2013. Where do we go from here? New perspectives on the black box in SHRM research. *Journal of Management Studies*, 50: 1448-1480.

*Katou, A. A., Budhwar, P. S., & Patel, C. 2014. Content vs. process in the HRM-performance relationship: An empirical examination. *Human Resource Management*, 53: 527-544.

*Kehoe, R. R., & Wright, P. M. 2013. The impact of high-performance human resource practices on employees' attitudes and behaviors. *Journal of Management*, 39: 366-391.

Kozlowski, S. W. J., Chao, G. T., Grand, J. A., Braum, M. T., & Kuljanin, G. 2013. Advancing multilevel research design: Capturing the dynamics of emergence. *Organizational Research Methods*, 6: 581-615.

Kozlowski, S. W., & Klein, K. J. 2000. A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions*: 3-90. San Francisco: Jossey-Bass.

Kreft, I. G., & de Leeuw, J. 1998. *Introducing multilevel modeling*. London: Sage.

*Kroon, B., Van De Voorde, K., & Van Veldhoven, M. 2009. Cross-level effects of high-performance work prac-tices on burnout: Two counteracting mediating mechanisms compared. *Personnel Review*, 38: 509-525.

Kuhn, T. S. 1962. *The structure of scientific revolutions*. Chicago: University of Chicago Press.

Lengnick-Hall, M. L., Lengnick-Hall, C. A., Andrade, L. S., & Drake, B. 2009. SHRM: The evolution of the field. *Human Resource Management Review*, 19: 64-85.

Lepak, D., & Boswell, W. 2012. Strategic HRM and employee organizational relationship (EOR). In L. M. Shore, J. A.-M. Coyle-Shapiro, & L. E. Tetrick (Eds.), *Essays in employee-organization relationships: Applications for the 21st century*: 455-483. New York: Psychology Press.

Lepak, D. P., Jiang, K., Han, K., Castellano, W. G., & Hu, J. 2012. Strategic HRM moving forward: What can we learn from micro perspectives? In G. P. Hodgkinson & J. K. Ford (Eds.), *International review of industrial and organizational psychology*, vol. 27: 231-259. Chichester, England: John Wiley & Sons.

Lepak, D. P., & Snell, S. A. 2002. Examining the HR architecture: The relationships among human capital, employ-ment, and HR configurations. *Journal of Management*, 28: 517-543.

*Li, X., & Frenkel, S. in press. Where *hukou* status matters: Analyzing the linkage between supervisor perceptions of HR practices and employee work engagement. *The International Journal of Human Resource Management*. doi:10.1080/09585192.2015.1137613

*Liao, H., & Chuang, A. 2004. A multilevel investigation of factors influencing employee service performance and customer outcomes. *Academy of Management Journal*, 47: 41-58.

*Liao, H., Toya, K., Lepak, D. P., & Hong, Y. 2009. Do they see eye to eye? Management and employee perspec-tives of HPWS and influence processes on service quality. *Journal of Applied Psychology*, 94: 371-391.

Mathieu, J. E., & Chen, G. 2010. The etiology of the multilevel paradigm in management research. *Journal of Management*, 37: 610-641.

Mathieu, J. E., & Taylor, S. R. 2007. A framework for testing meso-mediational relationships in organizational behavior. *Journal of Organizational Behavior*, 28: 141-172.

*Messersmith, J., Patel, P., Lepak, D., & Gould-Williams, J. 2011. Unlocking the black box: Exploring the link between HPWS and performance. *Journal of Applied Psychology*, 96: 1105-1118.

Molloy, J., Ployhart, R., & Wright, P. 2011. The myth of "the" micro-macro divide: Bridging disciplinary and system-level divides. *Journal of Management*, 37: 581-609.

Muthén, L. K. 2007. Mplus: Statistical analysis with latent variables (Version 4.21) [Computer software]. Los Angeles: Muthén & Muthén.

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

*Ogbonnaya, C., & Valizade, D. in press. High performance work practices, employee outcomes and organizational performance: A 2-1-2 multilevel mediation analysis. *The International Journal of Human Resource Management*. doi:10.1080/09585192.2016.1146320

*Oppenauer, V., & Van De Voorde, K. in press. Exploring the relationships between high involvement work system practices, work demands and emotional exhaustion: A multi-level study. *The International Journal of Human Resource Management*. doi:10.1080/09585192.2016.1146321

*Orlitzky, M., & Frenkel, S. J. 2005. Alternative pathways to high-performance workplaces. *The International Journal of Human Resource Management*, 16: 1325-1348.

Ostroff, C., & Bowen, D. E. 2000. Moving HR to a higher level: HR practices and organizational effectiveness. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations*: 211-266. San Francisco: Jossey-Bass.

Patel, P., Messersmith, J., & Lepak, D. 2013. Walking the tightrope: An assessment of the relationship between HPWS and organizational ambidexterity. *Academy of Management Journal*, 56: 1420-1442.

Peccei, R. 2004. *Human resource management and the search for the happy workplace* [Inaugural address]. Rotterdam, Netherlands: Erasmus Research Institute of Management.

Peccei, R., & Van De Voorde, K. 2014. Human resource management and performance. In D. E. Guest & D. Needle (Eds.), *Wiley encyclopedia of management: Vol. 5—Human resource management* (3rd ed.). Chichester, England: John Wiley & Sons.

Peccei, R., Van De Voorde, K., & Van Veldhoven, M. 2013. HRM, well-being and performance: A theoretical and empirical review. In J. Paauwe, D. Guest, & P. Wright (Eds.), *HRM & performance: Achievements and challenges*: 15-46. Oxford, England: Wiley-Blackwell.

Pfeffer, J. 1998. *The human equation: Building profits by putting people first*. Boston: Harvard Business Press.

*Piening, E. P., Baluch, A. M., & Salge, T. O. 2013. The relationship between employees' perceptions of human resource systems and organizational performance: Examining mediating mechanisms and temporal dynamics. *Journal of Applied Psychology*, 98: 926-947.

Preacher, K. J., Rucker, D. D., & Hayes, A. F. 2007. Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research*, 42: 185-227.

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

*Raineri, A. in press. Linking human resources practices with performance: The simultaneous mediation of collective affective commitment and human capital. *The International Journal of Human Resource Management*. doi:10.1080/09585192.2016.1155163

Raudenbush, S. W., & Bryk, A. S. 2002. *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage.

Salancik, G. R., & Pfeffer, J. 1978. A social information processing approach to job attitudes and task design. *Administrative Science Quarterly*, 23: 224-253.

Schneider, B. 1987. The people make the place. *Personnel Psychology*, 40: 437-453.

Shen, J. in press. Principles and applications of multilevel modeling in HRM research. *Human Resource Management*. doi:10.1002/hrm.21666

*Shen, J., Benson, J., & Huang, B. 2014. HPWS and teachers' work performance: The mediating role of quality of working life. *Human Resource Management*, 53: 817-833.

*Shin, S. J., Jeong, I., & Bae, J. in press. Do high-involvement HRM practices matter for worker creativity? A cross-level approach. *The International Journal of Human Resource Management*. doi:10.1080/09585192.2015.1137612

*Snape, E., & Redman, T. 2010. HRM practices, organizational citizenship behaviour, and performance: A multilevel analysis. *Journal of Management Studies*, 47: 1219-1247.

Snijders, T., & Bosker, R. 1999. *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. London: Sage.

Subramony, M. 2009. A meta-analytic investigation of the relationship between HRM bundles and firm performance. *Human Resource Management*, 48: 745-768.

*Takeuchi, R., Chen, G., & Lepak, D. P. 2009. Through the looking glass of a social system: Cross-level effects of HPWS on employees' attitudes. *Personnel Psychology*, 62: 1-29.

*Takeuchi, R., Lepak, D. P., Wang, H., & Takeuchi, K. 2007. An empirical examination of the mechanisms medi-ating between high-performance work systems and the performance of Japanese organizations. *Journal of Applied Psychology*, 92: 1069-1083.

Tay, L., Woo, S. E., & Vermunt, J. K. 2014. A conceptual and methodological framework for psychometric isomor-phism validation of multilevel construct measures. *Organizational Research Methods*, 17: 77-106.

Toh, S. M., Morgeson, F. P., & Campion, M. A. 2008. HR configurations: Investigating fit with the organizational context. *Journal of Applied Psychology*, 93: 864-882.

*Van De Voorde, K., & Beijer, S. 2015. The role of employee HR attributions in the relationship between HPWS and employee outcomes. *Human Resource Management Journal*, 25(1): 62-78.

Van De Voorde, K., Paauwe, J., & Van Veldhoven, M. 2012. Employee well-being and the HRM–organizational performance relationship: A review of quantitative studies. *International Journal of Management Reviews*, 14: 391-407.

*Van De Voorde, K., Veld, M., & Van Veldhoven, M. 2016. Connecting empowerment-focused HRM and labour productivity to work engagement: The mediating role of job demands and resources. *Human Resource Management Journal*, 26(2): 192-210.

*Veld, M., Paauwe, J., & Boselie, P. 2010. HRM and strategic climates in hospitals: Does the message come across at the ward level? *Human Resource Management Journal*, 20(4): 339-356.

*Vermeeren, B. 2014. Variability in HRM implementation among line managers and its effect on performance: A 2-1-2 mediational multilevel approach. *The International Journal of Human Resource Management*, 25: 3039-3059.

Wall, T. D., & Wood, S. J. 2005. The romance of human resource management and business performance, and the case for big science. *Human Relations*, 58: 429-462.

Wallace, J. C., Edwards, B. D., Paul, J., Burke, M., Christian, M., & Eissa, G. 2016. Change the referent? A meta-analytic investigation of direct and referent-shift consensus models for organizational climate. *Journal of Management*, 42: 838-861.

Wei, L. Q., & Lau, C. M. 2010. High performance work systems and performance: The role of adaptive capability. *Human Relations*, 10: 1487-1511.

*Wei, Y. C., Han, T. S., & Hsu, I. C. 2010. High-performance HR practices and OCB: A cross-level investigation of a causal path. *The International Journal of Human Resource Management*, 21: 1631-1648.

Whitman, D., Van Rooy, D., & Viswesvaran, C. 2010. Satisfaction, citizenship behaviors, and performance in work units: A meta-analysis of collective construct relations. *Personnel Psychology*, 63: 41-81.

*Wood, S., Van Veldhoven, M., Croon, M., & de Menezes, L. M. 2012. Enriched job design, high involvement management and organizational performance: The mediating roles of job satisfaction and well-being. *Human Relations*, 65: 419-445.

Wright, P., & Boswell, W. 2002. Desegregating HRM: A review and synthesis of micro and macro human resource management research. *Journal of Management*, 28: 247-276.

Wright, P., & Gardner, T. 2003. The HR-firm performance relationship: Methodological and theoretical challenges. In P. Sparrow & A. Howard (Eds.), *The new workplace: A guide to the human impact of modern working practices*: 311-328. Chichester, England: Wiley.

*Wright, P., Gardner, T., & Moynihan, L. 2003. The impact of HR practices on the performance of business units. *Human Resource Management Journal*, 13(3): 21-36.

Wright, P., Gardner, T., Moynihan, L., & Allen, M. 2005. The relationship between HR practices and firm perfor-mance: Examining causal order. *Personnel Psychology*, 58: 409-446.

*Wu, P. C., & Chaturvedi, S. 2009. The role of procedural justice and power distance in the relationship between high-performance work systems and employee attitudes: A multilevel perspective. *Journal of Management*, 35: 1228-1247.

*Zhong, L., Wayne, S. J., & Liden, R. C. 2016. Job engagement, perceived organizational support, high-perfor-mance HR practices, and cultural value orientations: A cross-level investigation. *Journal of Organizational Behavior*, 37: 823-844.