PRINCIPLES AND APPLICATIONS OF MULTILEVEL MODELING IN HUMAN RESOURCE MANAGEMENT RESEARCH

JIE SHEN

Multilevel modeling is important for human resource management (HRM) research in that it often analyzes and interprets hierarchal data residing at more than one level of analysis. However, HRM research in general lags behind other disciplines, such as education, health, marketing, and psychology in the use of a multilevel analytical strategy. This article integrates the most recent literature into the theoretical and applied basics of multilevel modeling applicable to HRM research. A range of multilevel modeling issues have been discussed and they include statistical logic underpinning multilevel modeling, level conceptualization of variables, data aggregation, hypothesis tests, reporting mediation paths, and cross-level interactions. An empirical example concerning complex cross-level mediated moderation is presented that will suffice to illustrate the principles and the procedures for implementing a multilevel analytical strategy in HRM research. © 2015 Wiley Periodicals, Inc.

Keywords: human resource management (HRM), multilevel modeling (MLM), multilevel structural equation modeling (MSEM), research methods

Human resource management (HRM) research often involves hierarchal data from more than one level of analysis. Individual employees are nested in teams or departments that are entrenched within organizations. In turn, organizations are nested in industries embedded in larger environments, such as geographic regions, nations or economic or political blocks. HRM, as a subset of organizational policies, is a higher-level variable, as individuals within the same organization/unit share the same HRM policies and practices, but individuals in different organizations/units do not (Ostroff & Bowen, 2000). Employee attitudinal and behavioral responses to HRM policies and practices may be similar in the same organization and different in others due to contextual effects (Bliese & Hanges, 2004). Ignoring the inherent dependence of hierarchal data would result in deflated standard errors and inflated values of model fit or correlations (Rowe & Hill, 1998). This type of dependence in data structure is likely to lead to gross errors of prediction if using nonmultilevel modeling statistical approaches such as ordinary least squares (OLS) regression, designed to analyze the same level of data (Snijders & Bosker, 2012).

Correspondence to: Jie Shen, School of Management, University of South Australia Business School, City West Campus, Adelaide, South Australia, SA5001, Phone: +61 430355248, E-mail: jie.shen@unisa.edu.au

Human Resource Management, November–December 2016, Vol. 55, No. 6. Pp. 951–965

Published online in Wiley Online Library (wileyonlinelibrary.com).
DOI:10.1002/hrm.21666

Multilevel regression models account for variance among variables at different levels, handling sources of errors more rigorously than OLS, although parameter estimates are not substantially different (Rowe & Hill, 1998). Multilevel regression models have been variously named as hierarchical linear models (HLM) (Raudenbush & Bryk, 2002), mixed effects or mixed models (Littell, Milliken, Stroup, & Wolfinger, 1996), random coefficient models (Kreft & de Leeuw, 1998), and variance component models (Longford, 1986). Despite the different nomenclature, these models all simultaneously test relationships within a certain level and between or across hierarchical levels, allowing researchers to disentangle effects of between- and within-group variance on the dependent variable while using individual independent variables at the individual level, and group independent variables at the group level (Hofmann, 1997).

The perceived complexity and difficulties involved in multilevel design, data collection and analysis, and result interpretation and presentation may have hindered HRM researchers from conducting multilevel research.

Multilevel modeling (MLM) also offers advantages over alternative techniques including disaggregation and aggregation approaches in dealing with hierarchal data (Osborne, 2000). The disaggregation approach reduces upper-level variables to a lower level. Individuals in the same unit are assigned the same mean unit-related scores, which ignore between-group variations. Consequently, shared variance is no longer accounted for, and the assumption of independence of errors is violated, resulting in inflated correlations. The aggregation approach raises the lower-level variables to the higher hierarchal level, thereby ignoring individual differences. This approach results in research findings focused on higher hierarchal-level predictability, misrepresenting the relationships between variables (Hofmann, 1997; Raudenbush & Bryk, 2002). Multilevel modeling corrects these violations (Osborne, 2000).

It is important for HRM researchers to integrate macro and micro levels of analysis by simultaneously taking into account organizational effects and individual effects (Hofmann, Griffin, & Gavin, 2000; Ostroff & Bowen, 2000; Wei, Han, & Hsu, 2010; Whitener, 2001). Recently, a growing number of empirical studies have conceptualized and measured general HRM practices or high-performance work systems (HPWSs) or high-performance HRM at higher levels (Aryee, Walumbwa, Seidu, & Otaye, 2012; Den Hartog, Boon, Verburg, & Croon, 2013; Jensen, Patel, & Messersmith, 2013; Liao & Chuang, 2004; Messersmith, Patel, & Lepak, 2011; Sun, Aryee, & Law, 2007; Takeuchi, Chen, & Lepak, 2009). By exploring cross-level relationships between HRM/HPWS and a range of employee outcomes, these studies contribute to HRM literature in bridging macro and micro perspectives. The use of multilevel modeling in education, health, psychology, organizational behavior, and marketing research began in the late 1970s (Mathieu & Chen, 2011; Raudenbush & Bryk, 2002). Despite the growing interest from HRM researchers, HRM research overall lags behind in the use of multilevel modeling. Currently, the majority of HRM research does not adopt multilevel theoretical perspectives, but is conducted at the same level of analysis. This indicates a great need for further multilevel HRM research, metaphorically described by Kulik (2012) as "picking high-hanging fruits" and "climbing the higher mountain" (p. 447).

A reason for the inadequate implementation of multilevel modeling in HRM research is a lack of guiding theoretical frameworks linking HRM practices and other variables across multiple levels (Ostroff & Bowen, 2000). Ostroff and Bowen's (2000) meso paradigm integrates individual-level, organization-level, and cross-level effects, contributing to filling this important literature gap and calling for more multilevel HRM research (Hofmann et al., 2000). Another general reason for such deficiency is that multilevel models have only recently become available, and computer software has become technically feasible for only a couple of decades. Multilevel analytical techniques primarily emerged from educational research and have only been recently introduced to HRM research. The perceived complexity and difficulties involved in multilevel design, data collection and analysis, and result interpretation and presentation may have hindered HRM researchers from conducting multilevel research. Therefore, it is helpful if theories and procedures for multilevel modeling are explained in a systematic manner using empirical HRM examples that HRM researchers new to multilevel modeling can easily follow without frustration.

The main purposes of this study are to distill the literature with reference to statistical logic and the most recent debates and development regarding multilevel modeling, and to demonstrate the main procedures for conducting multilevel HRM research. The procedures covered in this study include level conceptualization of variables, data handling (e.g., collection, preparation, and aggregation), preconditions for conducting multilevel analysis, hypothesis testing (direct effect, mediation, and moderation), and reporting results in statistical or graphical forms.

Published pedagogical multilevel modeling manuals often use the simplest model containing Level 1 direct effect, Level 2 direct effect, and cross-level interaction as examples. However, HRM models are commonly complex, including mediation, mediated moderation, or moderated mediation. To appropriately guide HRM researchers, this article helpfully presents an empirical example of cross-level mediated moderation to illustrate the procedure for performing complex multilevel analysis in HRM research. As such, this article equips HRM researchers with the most up-to-date multilevel modeling knowledge, contributing to the promotion of further multilevel HRM research.

Statistical Logic Underpinning Multilevel Modeling

Essentially, multilevel modeling investigates simultaneously within-unit and between-unit relationships by estimating within-unit and between-unit models separately (Osborne, 2000). For the sake of simplicity and clarity, the discussions throughout this paper focus on two-level models. The Level 1 model illustrates the relationships between individual level predictor variables and the individual level outcome variable. The outcome of estimating the Level 1 model is intercepts and slopes that vary between units because each unit has its own intercept and slope. The Level 2 model indicates how Level 2 predictor variables predict the varying intercepts and slopes resulting from the analysis of Level 1 model (Hofmann, 1997; Hofmann et al., 2000). As Level 1 regression parameters are regressed onto Level 2 variables in Level 2 analysis, a multilevel modeling approach is actually a regression of regressions (Arnold, 1992). The Level 1 model that does not contain Level 2 predictor variables is shown in Equation 1:

$$Y_{ij} = \beta_{oj} + \beta_{ij} X_{ij} + e_{ij} \quad (1)$$

where Y_{ij} = value of Level 1 dependent variable (DV) for individual i in unit j, β_{oj} = the intercept (average DV) for unit j, β_{ij} = regression coefficient slope associated with Level 1 predictor variable for unit j, X_{ij} = value of Level 1 predictor variable for individual i in unit j, and e_{ij} (i.e., r_{ij} or s^2) = the residual error at Level 1 nested with Level 2. The Level 2 model takes intercepts and slopes for units as dependent variables and uses Level 2 predictor variable as a covariate at Level 2. For the intercepts, the equation is shown in Equation 2 and for the slopes in Equation 3. For the intercepts:

$$\beta_{oj} = \gamma_{00} + \gamma_{01} Z_j + u_{0j} \quad (2)$$

where γ_{00} (gamma subzero zero) = Level 2 intercept term, which is the mean DV when controlling for Level 2 predictor variable, γ_{01} = Level 2 slope term, which is the mean effect of Level 2 predictor variable on DV scores, Z_j = Level 2 predictor variable of unit j, and u_{0j} = residual error when modeling variation in intercepts.

Equation 3 represents the main effect of Level 2 predictor variable Z on the between-unit variance in Level 1 dependent variable after controlling for Level 1 predictor variable. For the slopes:

$$B_{ij} = \gamma_{10} + \gamma_{11} Z_j + u_{ij} \quad (3)$$

where γ_{10} = Level 2 intercept term, which is the mean DV when controlling for Level-2 predictor variable; γ_{11} = Level 2 slope term, which is the mean effect of Level 2 predictor variable on DV scores; Z_j = Level 2 predictor variable of unit j; and u_{ij} (i.e., $t_{00,}$ tau subzero zero) = residual error at Level 2 when modeling variation in slopes.

Equation 3 represents a cross-level interaction in which Z moderates the relationship between X and Y. Integrating Equations 2 and 3 results in a mixed-effect model, as shown in Equation 4, which highlights the interaction term $\gamma_{10} X_{ij} Z_j$ as a cross-level moderation.

$$Y_{ij} = \gamma_{00} + \gamma_{01} Z_j + \gamma_{10} X_{ij} + \gamma_{10} X_{ij} Z_j + u_{0j} + u_{ij} X_{ij} + e_{ij} \quad (4)$$

Multilevel analyses generate both fixed effects at Level 1, as shown in the first four terms, with gammas γ in Equation 4 nested within Level 2, and random effects shown in the last three terms that vary across units (Hofmann et al., 2000). u_{0j} and u_{ij} are residual error terms.

To test the hypothesized relationships indicated in Equation 4, the following conditions must be met:

- There is systematic within- and between-group variance in the DV. Multilevel analysis partitions within- and between-group variance. Systematic within-and between-group variance in the DV indicates the DV is influenced by both individual factors and group level factors, satisfying the need for examining the effect of group level factors.
- Mean Level 1 slopes in the DV across organizations are significantly different from zero. This indicates significant variance in the DV is, at least partially, due to Level 1 predictor variable.
- There is significant variance in Level 1 intercepts (the average score of the DV across organizations). This indicates significant variance in the DV is due to Level 2 predictor variable.

- There is significant variance in Level 1 slopes, indicating that while significant variance in the DV is caused by Level 1 predictor variable, the DV also differs in groups by Level 2 predictor; showing Level 1 predictor variable and Level 2 predictor variable interactively influence the DV.

Variable Level Conceptualization

According to Rousseau (1985, p. 4), the "[l]evel of measurement refers to the entities from which the data are drawn or are attached (e.g., raters, individuals, organizations etc.)." Lower-level entities do not have to be individuals; they can be groups, departments, organizations or regions. Repeated measurements of individuals may also be examined (Luke, 2004). When research is conducted at an interpersonal or intrapersonal level of analysis, the person is the higher-level unit. Multilevel regression analysis can also be applied to longitudinal data, where levels are defined by the measurement occasions nested within individuals (Snijders, 1996; Willms & Raudenbush, 1989). With the traditional MLM approach, mostly due to the constraints of computer software, the outcome variable is always situated at the lowest level of the hierarchy (Castro, 2002; Hofmann, 1997; Hofmann et al., 2000; Luke, 2004). In other words, traditional multilevel modeling can be used only to explore the effects of higher-level variables on Level 1 variables or the effects of the same-level predictor variables on Level 1 outcome, rather than the effects of lower-level variables on higher-level outcomes. However, with multilevel structural equation modeling (MSEM), the outcome variable can be situated at higher levels. For example, if a Level 1 predictor variable influences a Level 2 outcome variable via a Level 1 mediating variable, the model is called a 1-1-2 design. If a Level 2 predictor variable influences a Level 2 outcome variable via a Level 1 mediating variable, the model is a 2-1-2 design (Preacher, Zyphur, & Zhang, 2010).

Traditional multilevel modeling can be used only to explore the effects of higher-level variables on Level 1 variables or the effects of the same-level predictor variables on Level 1 outcome, rather than the effects of lower-level variables on higher-level outcomes.

In HRM research, in addition to HRM practices, a wide range of variables such as concern for employee climate (Takeuchi et al., 2009), service climate (Liao & Chuang, 2004) and empowerment climate (Aryee et al., 2012; Den Hartog et al., 2013), organizational communication (Den Hartog et al., 2013), organizational performance, such as service performance and customer satisfaction (Liao & Chuang, 2004), and financial performance (Den Hartog et al., 2013) have been conceptualized and measured at higher levels, such as unit, branch, department, and organization. Employee work attitudes and behaviors traditionally deemed as individual-level constructs can also be aggregated to higher levels. For instance, job satisfaction, affective commitment, psychological empowerment, and organizational citizenship behavior have been aggregated to the departmental level in Messersmith et al. (2011) and organizational level in Sun et al. (2007). Whether to conceptualize and measure a variable at a higher level should be determined by research design and theoretical and methodological rationales.

Empirical Example

The illustrative research project is designed to explore the effect of organizational high commitment HRM practices on individual employee knowledge sharing behavior and the underlying mechanisms to answer "why" and "when" questions. Several studies have explored the relationship between HRM and knowledge sharing (Collins & Smith, 2006). However, previous research was conducted at the same level of analysis, not taking into account the interaction of organizational contextual influence and individual effects. Moreover, the mechanism via which high-commitment HRM influences employee knowledge sharing is unclear. The example project addresses these limitations. High-commitment HRM practices, such as developmental appraisals, comprehensive training and development, and competitive and equitable pay, create "conditions that encourage employees to identify with the goals of the organization and work hard to accomplish those goals" (Whitener, 2001, p. 517).

In their meso paradigm, Ostroff and Bowen (2000) noted cross-level effects of organization-level HRM practices on individual work attitudes and behaviors. Building on this meso paradigm, this article hypothesizes a positive cross-level effect of high-commitment HRM on employee knowledge sharing. High-commitment HRM is indicative of personified organizational support for employees, who will reciprocate organizational support with more positive discretional work behaviors, such as knowledge sharing. Hence, perceived organizational support (POS) would mediate the relationship between high-commitment HRM and knowledge sharing. As knowledge sharing is cooperative organizational citizenship behavior, the effect of HRM and POS on knowledge sharing is influenced by the organizational cooperative climate. That is, organizational cooperative climate moderates the relationships of high-commitment HRM and POS with knowledge sharing. Specifically, this article tests the following hypotheses:

Hypothesis 1: High-commitment HRM will be positively related to employee knowledge-sharing behavior.

Hypothesis 2: High-commitment HRM will influence employee knowledge-sharing behavior through the mediation of perceived organizational support.

Hypothesis 3a: Organizational cooperative climate will moderate the relationship between high-commitment HRM and employee knowledge-sharing behavior.

Hypothesis 3b: Organizational cooperative climate will moderate the relationship between perceived organizational support and employee knowledge-sharing behavior.

In this illustrative research, employee knowledge sharing is the individual-level dependent variable. High-commitment HRM is the organization-level predictor influencing employee knowledge sharing via the mediation of POS at the individual level. Organizational cooperative climate is conceptualized as the organizational-level variable moderating the direct effect of high-commitment HRM and the second stage of the mediation of POS. Figure 1 depicts the conceptual model indicating the multilevel nature of the research design: Hypothesis 1 proposes a cross-level main effect; Hypothesis 2 is a meso-mediation relationship with a 2-1-1 design; Hypotheses 3a and 3b are cross-level moderations. Jointly, the model is a meso-mediated moderation, indicating the need to use a multilevel modeling strategy. This example therefore will sufficiently demonstrate how to conduct complex multilevel HRM research.

Sample and Measures

A sample of 738 employees from 30 firms in the People's Republic of China was used. The data for the study were collected between July and October 2012 through the network of Chinese MBA students who also held managerial positions in the participating firms. These firms cover manufacturing, retail, finance, electronics, distillers, communication, food processors, chemicals, and hotels—representing various industry sectors. The mean number of employees is 532, with a standard deviation (SD) of 741. The multisourced data were used to avoid common method variance (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). Specifically, employees answered questions about high-commitment HRM, POS, and cooperative climate. Supervisors rated their subordinates' knowledge-sharing behavior. Each firm was required to randomly select 50 employees from one department or production/service unit to participate in the study. The total number of received matched completed surveys was 761, accounting for a 51 percent individual-level response rate. The number of usable surveys is 738.

The measure for high-commitment HRM practices was adapted from the five items developed by Snell and Dean (1992). A sample item is "My organization provides comprehensive, adequate training to employees." POS was measured using the eight-item scale adapted from Eisenberger, Huntington, Hutchison, and Sowa (1986). A sample item is "My organization cares about my opinions." Cooperative climate was measured using a five-item scale adapted from Chatman and Flynn (2001). A sample item is "There is harmony within my team." Knowledge-sharing behavior was measured using seven items adapted from Bock and Kim (2002). A sample item is "He or she always provides constructive ideas to help colleagues improve performance." Likert-type scales ranging from 1 = strongly disagree to 5 = strongly

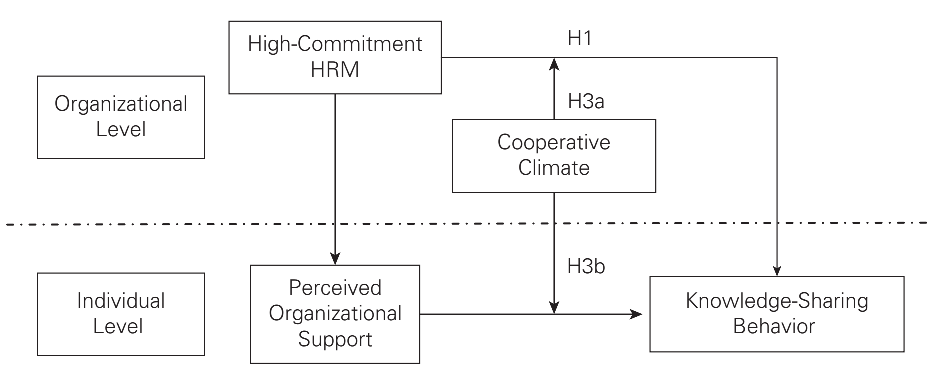

FIGURE 1. The Conceptual Model and Hypotheses

agree were used for all study variables. The scales have Cronbach alpha values between .84 and .89. Gender, age, education, hierarchal position, and tenure are controlled at the individual level, and organizational size (log of employee numbers) and response rate at the organizational level.

Data Issues

Data Requirements

A sufficient sample size for multilevel modeling is important for accurate estimation in terms of regression coefficients and variance. Statistical power for multilevel models depends on the number of both individual observations at the individual level and groups at the unit level, as Level 1 sample size influences the statistical power to detect Level 1 direct effects and Level 2 sample size is relevant to the statistical power to detect Level 2 direct effects (Raudenbush & Liu, 2000). The requirement for both large individual sample size and group sample size is regarded as a disadvantage of multilevel modeling (Kreft & de Leeuw, 1998). However, "the group-level sample size is generally more important than the total sample size or observations per group, with a large individual level sample size partially compensating for a small number of groups" (Maas & Hox, 2006, p. 87).

Statistical power for multilevel models depends on the number of both individual observations at the individual level and groups at the unit level, as Level 1 sample size influences the statistical power to detect Level 1 direct effects and Level 2 sample size is relevant to the statistical power to detect Level 2 direct effects.

Specifically, according to Kreft and de Leeuw (1998), a sample of 20 is recommended as the smallest acceptable number for groups, and five for individual observations per group. To have sufficient power (e.g., .90) to detect cross-level interactions, a sample of 30 units containing 30 individuals each is desirable (Hofmann et al., 2000). Notably, in general, the larger the group-level sample, the stronger the power of predictability (Hofmann, 1997). Hence, it would be helpful to increase sample size, the group-level sample size in particular, especially when researchers are interested in effects of higher-level predictor variables.[1] The illustrative sample contains 30 firms and the mean number of responses at Level 1 is 24.6; meeting the recommended requirements for sample size for multilevel analysis. As for missing data, multilevel modeling handles missing data only at Level 1. Missing data are not allowed at Level 2 or above. For this reason, groups with missing data at Level 2 or above should be removed.

Data Aggregation

Some computer programs such as HLM require separate datasets for performing multilevel analyses. In the case of a two-level analysis, two datasets are required—one containing individual-level variables and the other Level 2 variables. The group ID links two datasets. Statistical Package for the Social Sciences (SPSS) has the function to aggregate data: "If global measures of constructs are not available, data must be gathered from multiple employees within a firm, and from multiple firms" (Ostroff & Bowen, 2000, p. 253).

Although HRM practices should be conceptualized at the organizational level, it is preferable if the data are collected from individual employees due to the importance of employee perceptions of HRM practices (Ostroff & Bowen, 2000). Liao, Toya, Lepak, and Hong (2009) noted significant differences between managerial perceptions and employee perceptions of HRM practices in the same organization. In the illustrative project, the data for high-commitment HRM and cooperative climate were collected from individual employees. When higher-level variables are composites of lower-level variables, researchers should justify the aggregation of the data for lower-level variables (Mathieu & Chen, 2011). The following two preconditions should be met in order to aggregate data as Level 2 variables (Rousseau, 1985).

Condition 1: A high level of interrater agreement for Level 2 variables within the Level 2 units.

Interrater agreement or homogeneity means the reliability of unit-level variables takes into account differences within units relative to differences between units. R_{wg} has been developed by Bliese (2000) and James, Demaree, and Wolf (1984) to assess the level of interrater agreement. R_{wg} can be calculated using the following equation (LeBreton & Senter, 2008, p. 819):

$$R_{wg(J)} = J[1-(\bar{s}^2_{Xj})/\sigma^2_E] \,/\, J[1-(\bar{s}^2_{Xj})/\sigma^2_E)+\bar{s}^2_{Xj})/\sigma^2_E] \qquad (5)$$

where J = the number of items ranging from $j = 1$ to J, X = an observed score, typically measured on an interval scale of measurement, $\bar{s}^2_{Xj}$ = the mean observed variance on X, σ^2_E = the variance expected when there is a complete lack of agreement among the raters and $\bar{s}^2_{Xj})/\sigma^2_E$ = the proportion of error variance caused by random responding.

Readers are referred to LeBreton and Senter (2008, pp. 841–845) for the syntax for calculating

r_{wg} using SPSS. According to James et al. (1984), an r_{wg} greater than .70 is acceptable and the higher the value of r_{wg}, the stronger the within-group agreement of the construct. R_{wg} is assessed in one group at a time. If r_{wg} is below .70 for some groups, researchers should determine how many groups and variables have low r_{wg} values and why. It is common to report either the range or average for each group for each variable (LeBreton & Senter, 2008). With the illustrative sample, r_{wg}s range from .74 to .91 for high commitment HRM, and from .75 to .94 for cooperative climate, respectively.

Condition 2: Systematic between-group variations in Level-2 variables.

Generally, researchers should adopt the following three procedures to investigate between-group variations in Level 2 variables. Firstly, one-way ANOVA (analysis of variance) is performed to examine between-group variations. ANOVA at this stage can be performed on SPSS. Second, intraclass correlation (ICC(1)) is calculated to check the proportion of variance due to team variability using the following equation:

$$\text{ICC(1)} = [\text{MSB} - \text{MSW}] / [\text{MSB} + (n - 1) \times \text{MSW}] \quad (6)$$

where MSB is the between-group mean square, MSW is the mean square within group and n is the average number of members within groups. MSB and MSW can be obtained by conducting one-way ANOVA on SPSS in which Level 2 variable is the DV and the group ID is IV. Bliese (2000) suggests that ICC(1) values different from zero are desirable, with values close to .20 indicating high scores for group-level analysis.

Thirdly, reliability of the mean (ICC(2)) is calculated to examine the extent to which teams can be used to reliably differentiate in terms of individuals' ratings. ICC(2) can be calculated using Equation 7. ICC(2) values greater than .60 are regarded as desirable (Glick, 1985).

$$\text{ICC(2)} = (\text{MSR} - \text{MSW}) / \text{MSR} \quad (7)$$

With the illustrative sample, one-way ANOVA analyses demonstrate significant variations in high commitment: SRHRM ($F(29) = 3.41$, $p < .001$) and in cooperative climate ($F(29) = 2.73$, $p < .01$) among the 30 participating companies. ICC(1) and ICC(2) are .18 and .64 for high-commitment HRM, and .16 and .71 for cooperative climate, respectively. These results show a high level of interrater agreement for Level 2 variables within the Level 2 units and systematic between-group variations in Level 2 variables, justifying the aggregation of high-commitment HRM and cooperative climate as Level 2 constructs.

Data Centering

Centering of predictor variables reduces unnecessary multicollinearity and improves the interpretability of lower-order coefficients in multilevel analysis (Hofmann & Gavin, 1998; Kreft & de Leeuw, 1998; Mathieu, Aguinis, Culpepper, & Chen, 2012). Group mean centering centers predictors on group means. In this case, the intercept is the average outcome for each group; allowing interpretation of parameter estimates as person-level effects within each group. In the grand mean centered model, predictors are centered on overall means. This represents the group mean value for a person with a (grand) average on every predictor. It is recommended that Level 1 predictors normally are group mean centered to more accurately estimate intercepts. Level 2 predictors are suggested to be grand mean centered. This results in Level 2 intercept being equal to the mean score of the outcome variable (Mathieu et al., 2012). Dummy variables can be centered, although dummy variables do not change the interpretation of the intercepts when group mean-centering is employed (Hofmann & Gavin, 1998; Kreft & de Leeuw, 1998).

Multilevel analysis involves testing four hierarchical models: null model, random intercepts model, intercepts-as-outcome model, and slopes-as-outcomes model.

When interaction terms are created using the variables at the same level, researchers need to create interaction terms separately (e.g., on SPSS) and normally use grand mean centering for both focal predictors and moderators (Mathieu et al., 2012). Preacher et al. (2010) suggested that if more than a couple of the variables have ICCs below .05, the estimation of the indirect effect is likely to be unstable with a potentially large bias. Under this circumstance, group mean should be used to obtain stable indirect effect. Accordingly, in the illustrative research, individual-level control variables are not centered, Level 1 variables are group mean centered, and Level 2 variables are grand mean centered.

Hypothesis Testing

Multilevel analysis involves testing four hierarchical models: null model, random intercepts model, intercepts-as-outcome model, and slopes-as-outcomes model. A null model is also labeled *an unconstrained or intercept-only model,* which includes no explanatory variables. A null model allows intercepts to vary and assumes that slopes are fixed across higher-level units. It is used to estimate whether the residual variance in the individual-level model by Level 2 units is significantly different from zero, and to confirm whether multilevel modeling is necessary. A random intercepts model is also labeled *random coefficient regression model*. This model tests the relationship between Level 1 predictor variable and the same-level outcome variable. An intercepts-as-outcomes model is also labeled a *fixed slope model or means-as-outcomes model*. It is used to examine the direct effect of the higher-level predictor variable on the lower level criterion variable. A slopes-as-outcomes model is also labeled *random slopes model,* which examines whether cross-level interactions significantly account for some variance in Level 1 slopes.

It is prudent to note that to test a conventional simple multilevel model containing Level 1 direct effect, Level 2 direct effect and cross-level interaction one normally tests these four hierarchical models in sequence as shown earlier by entering the control variables, the predictor variables, and interaction terms at different stages. However, some research does not need to test all four models. For example, one does not need to test the random intercepts model if the researcher is interested only in direct cross-level effects. Moreover, the four models do not always have to be tested sequentially and can sometimes be tested simultaneously. For instance, if the research interest is cross-level mediation, the random intercepts model and intercepts-as-outcome model should be estimated simultaneously (Zheng, Zyphur, & Preacher, 2009). A slopes-as-outcomes model is often tested together with a random intercepts model. When tested together, the combined model is labeled a random intercepts and slopes model.

In the illustrative research, our first research objective is to explore Level 2 direct effect. Hence, we first estimate a null model, which is followed by an intercepts-as-outcomes model. Due to the fact that our second research objective is to explore cross-level mediation, our next step is to simultaneously estimate a random intercepts model and an intercepts-as-outcome model by adopting the MSEM approch. Finally, we estimaite a slopes-as-outcomes model to test cross-level moderation and an intercepts-as-outcomes model to test Level-1 moderation. For the purpose of the illustrative analysis, it is assumed that all assumptions of multilevel modeling are adequately met. Specifically, there is no multicollinearity, normal distribution of error terms at every level of the model, homogeneity of variance and independence of observations (see Hofmann et al., 2000, for the details of assumptions of MLM). Firms, groups, and units are treated as interchangeable throughout the article when referring to Level 2 units.

Null Model

In the example, the null model is tested using the following equation:

$$\textit{Knowledge sharing behavior} = \gamma_{00} + u_{0j} + r_{ij} \qquad (8)$$

where the fixed part of the model γ_{00} = grand mean knowledge sharing (i.e., the average score of employee knowledge sharing across all firms), the random part of the model u_{0j} = individual (within-group) variance in knowledge sharing, and r_{ij} = between-group variance in knowledge sharing. As there are no explanatory variables in the model, the variance is used in the calculation of the ratio of the between-group variance to the total variance, termed ICC for the criterion variable. In the illustrative example, the result of the analysis using HLM 7 is statistically significant [$u_{0,}$ $\chi^2(29) = 18.\ 81$, $p < .001$]. Hence, additional variance exists by organizational level predictors, justifying the application of multilevel modeling. ICC is calculated using Equation 9:

$$\text{ICC} = \tau_{00} / (\tau_{00} + \sigma^2(r_{ij})) \qquad (9)$$

ICC = .07941 / (.07941 + .22995) = .26, indicating 26% of variance in Level 1 criterion variable resides at a higher level of analysis (i.e., between firms). This result further suggests the implementation of a multilevel data-analytic strategy (Snijders & Bosker, 2012).

Intercepts-as-Outcomes Model

An intercept term is generally interpreted as the expected average of the dependent variable given that all predictors in a model are equal to zero. In the example, the variance in the average score of employee knowledge sharing across organizations may be due to organizational factors such as high-commitment HRM. This model is run to estimate the following equation (Equation 10) that also includes the control variables to confirm Hypothesis 1:

$$\textit{Knowledge sharing behavior} = \gamma_{00} + \gamma_{01}\,(\text{HRM}) + \gamma_{10}\,(\text{gender}) + \gamma_{20}\,(\text{age}) + \gamma_{30}\,(\text{position}) + \gamma_{40}\,(\text{education}) + \gamma_{50}\,(\text{tenure}) + \gamma_{60}\,(\text{POS}) + u_0 + r_{ij} \qquad (10)$$

where γ_{00} = Level 2 intercept, γ_{01} = Level 2 slope (Hypothesis 1), γ_{10}–γ_{60} = mean (pooled) slopes, u_0 = residual intercept variance, and r_{ij} = Level 1 residual variance.

The analysis using HLM 7 shows high commitment HRM is significantly and positively related to employee knowledge sharing behavior (γ_{01} = .35, t = 4.1, p < .001) holding POS constant. Every SD increase in high-commitment HRM at the organization-level results in .35 SD increase in employee knowledge sharing behavior. Hence, Hypothesis 1 receives support.

Random Intercepts Model

As this model tests the relationship between Level 1 predictor variable and the outcome variable, with a simple multilevel model without cross-level mediation, it requires testing prior to testing the intercepts-as-outcomes model. In the illustrative research, Hypothesis 2 proposes cross-level mediation, which involves testing a random intercepts model, and an intercepts-as-outcomes model (Equation 10). The random intercepts model is tested using Equation 11:

$$\textit{Knowledge sharing behavior} = \gamma_{00} + \gamma_{10}\,(\text{gender}) + \gamma_{20}\,(\text{age}) + \gamma_{30}\,(\text{position}) + \gamma_{40}\,(\text{education}) + \gamma_{50}\,(\text{tenure}) + \gamma_{60}\,(\text{POS}) + u_0 + r_{ij} \quad (11)$$

where γ_{00} = the mean of intercepts across groups, γ_{10}–γ_{60} = means of the slopes across groups, u_0 = variance in intercepts, and r_{ij} = Level 1 residual variance. A simple multilevel model does not involve the procedure for estimating indirect effect, as demonstrated next.

Test for Cross-Level Mediation

Hypothesis 2 in the illustrative research is cross-level mediation that requires testing for cross-level indirect effect. Traditional MLM tests cross-level indirect effect adopting Baron and Kenny's (1986) approach developed for testing single-level mediation. For example, Freedman and Schatzkin (1992) have suggested that the meso-mediation effect can be calculated using the formula: "$\gamma_{01\ \textit{without controlling for M (mediator)}} - \gamma_{01\ \textit{controlling for M}}$" or "$\gamma_{01\ \textit{without controlling for M (mediator)}} \times \gamma_{01\ \textit{controlling for M}}$." Recently, some authors, including Kenny (Kenny, Kashy, & Bolger, 1998), have recognized that the existence of mediation does not require the existence of a significant direct relationship between the independent variable and the dependent variable.

Accordingly, to test Hypothesis 2, calculating the cross-level indirect effect that can be obtained using the coefficient between high commitment HRM and POS ($\gamma_{01,}$ in Equation 12), the coefficient between POS and knowledge sharing ($\gamma_{60,}$ in Equation 11) is needed:

$$\textit{POS} = \gamma_{00} + \gamma_{01}\,(\text{HRM}) + \gamma_{10}\,(\text{gender}) + \gamma_{20}\,(\text{age}) + \gamma_{30}\,(\text{position}) + \gamma_{40}\,(\text{education}) + \gamma_{50}\,(\text{tenure}) + u_0 + r_{ij} \quad (12)$$

Nevertheless, this procedure has been criticized by authors such as Zheng et al. (2009) and Preacher et al. (2010) for bias of indirect effects, not taking into account measurement error, providing goodness-of-fit indices and the inability to model effects involving higher level dependent variables. To address these limitations, Preacher et al. (2010) and Zheng et al. (2009) suggest using MSEM, which has a range of advantages such as allowing for separate estimation of between-group and within-group relationships, simultaneous estimation first stage and second stage mediation, and treating variables as latent. Preacher has provided syntax for performing MSEM with the bootstrapping procedure to test cross-level mediation using Mplus software.[2] According to Preacher et al. (2010), cross-level indirect effects using MSEM approach are more accurate than using traditional MLM approach. The procedure for using the MSEM approach to test Hypothesis 2 is demonstrated later.

A partial mediation model was fitted with the direct relationship between high commitment HRM and knowledge sharing, as well as the indirect relationship via POS and compared with a full mediation model only with the indirect relationship: "Any mediation effect in a model at least one of X, M, or Y is assessed at Level-2 must occur strictly at the between-group level" (Preacher et al., 2010, p. 210). Hence, both models are assessed at the between-group level.

It is important to note that there are yet no standard cutoffs for fit indices for multilevel modeling.

It is important to note that there are yet no standard cutoffs for fit indices for multilevel modeling. Hence, this article refers to the fit indices for single level SEM models. The partial mediation model fits well into the data: $\chi^2(9)$ = 17.52, the comparative fit index (CFI) = 1.00, the Tucker Lewis (TLI) = 1.00, the root mean square error of approximation (RMSEA) = .05, the standardized root mean square residual (SRMR) within = .01, and the SRMR between = .04. This partial mediation fits significantly better [$\Delta\chi^2_{(1)}$ = 6.22, p < .05] than the full mediation model [$\chi^2(8)$ = 23.74, CFI = .87, TLI = .86, RMSEA = .06, SRMR within = .03, and SRMR between = .18]. Hence, the partial mediation model is deemed as the final model.

Checking coefficients at the within-organization level, POS is significantly related to knowledge sharing behavior (γ = .22, p < .01). At the between-organization level, high-commitment HRM is significantly related to POS (γ = .77, p < .001) and knowledge sharing (γ = .40, p < .001). POS is related to knowledge sharing (γ = .32, p < .001). The indirect effect at the between-organization level from high commitment via POS to knowledge sharing is significant (.25, $p < .01$). The bootstrapping test shows 95 percent of the confidence interval of the indirect effect is [.39, .10] and does not contain zero.

The bootstrapping procedure (Preacher & Hayes, 2008) and the PRODCLIN procedure (MacKinnon, Fritz, Williams, & Lockwood, 2007) are currently regarded as effective in rigorously testing the significance of indirect effects. It is argued that the Sobel test does not work well in small samples and is not recommended for use if researchers have the assess to raw data (Preacher & Hayes, 2008). The bootstrapping approach does not impose distributional assumptions resulting in greater control of Type I error rates and higher power; it provides a better alternative with a small sample size (MacKinnon et al., 2007; Preacher & Hayes, 2008).

Researchers must take the following two steps to test interactions with multilevel modeling: Step 1—testing whether cross-level moderation exists; and Step 2—testing significance of slopes of interactions.

Pseudo R^2 for the whole model and for the variance attributed to high commitment HRM can be calculated. Pseudo R^2 for the whole model is estimated using Equation 13 and for variance by high-commitment HRM Equation 14, respectively.

Pseudo R^2 for the model:

$$(\sigma^2_{\text{null}} - \sigma^2_{\text{intercepts-as-outcomes}}) \,/\, \sigma^2_{\text{null}} \qquad (13)$$

Pseudo R^2 for HRM:

$$(\tau_{00\ \text{random intercept}} - \tau_{00\ \text{intercepts-as-outcomes}}) \,/\, \tau_{00\ \text{random intercept}} \qquad (14)$$

Results show that the pseudo R^2 for the model is (.23 – .18) / .23 = .22 and for high-commitment HRM is (.014 – .012)/.014 = .14, indicating the whole mediation model and high-commitment HRM explain 22 percent and 14 percent of variance in Level 1 criterion variable, respectively. Combined, Hypothesis 2 is confirmed.

Slopes-as-Outcomes Model

Interactions can occur between two predictors at Level 1, at Level 2, or cross-level. Strictly speaking, only cross-level interaction is a slopes-as-outcomes model. The interaction between two predictors within Level 2 is a means-as-outcomes model. In the illustrative research, Hypothesis 3a is a slopes-as-outcomes model while Hypothesis 3b is a means-as-outcomes model.

Researchers must take the following two steps to test interactions with multilevel modeling: Step 1—testing whether cross-level moderation exists; and Step 2—testing significance of slopes of interactions.

Step 1: Testing whether cross-level moderation exists. Although Edwards and Lambert's (2007) procedure was initially developed for single level mediated moderation, it is equally applicable to multilevel mediated moderation and has been used previously by Liu, Liao, & Loi (2012) and Tse, Dasborough, and Ashkanasy (2008). This procedure was adopted by bringing moderations on the direct effect of the independent variable and in the second stage of mediation into the equation, as demonstrated in Equation 15:

$$\textit{Knowledge sharing behavior} = \gamma_{00} + \gamma_{01}\,(\text{HRM}) + \gamma_{02}\,(\text{cooperative climate}) + \gamma_{03}\,(\text{HRM} \times \text{cooperative climate}) + \gamma_{10}\,(\text{gender}) + \gamma_{20}\,(\text{age}) + \gamma_{30}\,(\text{position}) + \gamma_{40}\,(\text{education}) + \gamma_{50}\,(\text{tenure}) + \gamma_{60}\,(\text{POS}) + \gamma_{61}\,(\text{cooperative climate} \times \text{POS}) + u_0 + u_5\,(\text{POS}) + r_{ij} \qquad (15)$$

where g_{00} = Level 2 intercept, g_{01}, g_{02}, g_{03}, and g_{61} = Level 2 slopes, g_{10}, g_{20}, g_{30}, g_{40}, g_{50} and g_{60} = Level 2 intercepts, u_0 = residual intercept variance, u_5 = residual slope variance and *rij* = Level 1 residual variance. The results show that the interaction of high commitment HRM and cooperative climate is significantly related to knowledge sharing: $\gamma_{03} = .09$, $t = 4.1$, $p < .05$, and the interaction of POS and cooperative climate is significantly related to knowledge sharing: $\gamma_{61} = .07$, $t = 4.7$, $p < .05$. Notably, although γ_{02}, the main effect of cooperative climate on knowledge sharing is nonsignificant, the significance of the interaction terms still requires testing.

Following the procedure suggested by Bliese (2002), Equation 16 was used to calculate a pseudo R^2 for the whole model (.23 – .13) / .23 = .43 and Equation 17 for the moderator variable, cooperative climate (.012 – .009) / .012 = .25.

Pseudo R^2 for the model:

$$(\sigma^2_{\text{null}} - \sigma^2_{\text{intercepts-as-outcomes}}) \,/\, \sigma^2_{\text{null}} \qquad (16)$$

Pseudo R^2 for cooperative climate:

$$(\tau_{00\ \text{intercepts as outcomes}} - \tau_{00\ \text{slopes as outcomes}}) \,/\, \tau_{00\ \text{intercepts as outcomes}} \qquad (17)$$

Step 2: Testing significance of slopes of interactions. Preacher, Curran and Bauer (2006) have created an online calculator for simple slope tests.[3] Hypothesis 3a is the same as Preacher's Case 2 and Hypothesis 3b is the same as Preacher's Case 3. The values for fixed coefficients and the asymptotic covariance matrix of fixed regression estimates are needed to calculate simple slopes.

The values of the moderator, as suggested by Aiken and West (1991), may be one SD below the mean and one SD above the mean when the moderator is a continuous variable, or zero and one if it is dichotomous.

Using Preacher's calculator, the following results were obtained: $\gamma_{01} = .19$, $p < .01$ when cooperative climate is low and $\gamma_{01} = .51$, $p < .001$ when cooperative climate is high for high commitment HRM; $\gamma_{60} = .16$, $p < .05$ when cooperative climate is low and $\gamma_{60} = .47$, $p < .001$ when cooperative climate is high for POS at the between-organization level. At the individual level, $\gamma_{60} = .13$, $p < .05$ when cooperative climate is low and $\gamma_{60} = .37$, $p < .001$ when cooperative climate is high for POS. Taken together, when cooperative climate is low, the relationships of high commitment HRM and POS with knowledge sharing are weaker and when high, the relationships are stronger. Therefore, Hypotheses 3a and 3b are supported.

The results of the moderated path analyses are shown in Table I. Some statistical packages such as SAS, SPSS, R, and HLM, have graphing features for plotting cross-level interactions. Online utility for plotting HLM two- and three-way interactions are also available.[4]

Statistical Packages

Several major statistical packages, such as HLM, R, SPSS, MPlus, LISREL, and MLWiN are available for performing multilevel analysis. Each package has advantages and disadvantages in terms of operation and output production. MPlus is gaining popularity as it provides desired output for ICC and effects at different stages of mediation, including direct, first stage mediation, second stage mediation, indirect and total effects. It is also more suitable for performing MSEM, which produces model fit indices, and separates within- and between-group effects. However, Mplus does not tolerate errors in data or in syntax.

The HLM program is commonly utilized for conducting two- or three-level analysis. The advantages of HLM are the requirement of fewer assumptions to be met than other programs (Raudenbush & Bryk, 2002), accommodating a lack of sphericity, missing data, small and/or discrepant group sample sizes, and heterogeneity of variance across repeated measures (Osborne, 2000). More importantly, for novice multilevel researchers, it is user-friendly software. The major disadvantage is that it is unsuitable for performing MSEM. The procedure for conducting simple multilevel analysis without cross-level mediation using the HLM software including program setup and hypothesis tests is demonstrated in the Appendix.

Concluding Remarks

Multilevel modeling is important to HRM research as it takes into account the interactions of contextual factors of higher levels of units and individual factors, so that predicting accuracy increases. Despite recent growing interest in the use of multilevel analytical strategies, HRM research lags behind other disciplines such as education, marketing, and psychology, in this regard. This study aims to serve as a guide for a general readership of HRM researchers in understanding multilevel modeling concepts and procedures for conducting multilevel research.

This study aims to serve as a guide for a general readership of HRM researchers in understanding multilevel modeling concepts and procedures for conducting multilevel research.

This study discusses, from a theoretical perspective, the importance of multilevel modeling in HRM research and the advantages of multilevel

TABLE I Results of the Moderated Path Analyses

Cooperative Climate		HCHRM (X) POS (M) Knowledge Sharing (Y): First Stage PMX	Second Stage PYM	Direct Effect PYX	Indirect Effect PMX*PYM	Total Effect PYX+ PMX*PYM
Low levels of cooperative climate	Between-level	β = .77, $p < .001$	β = .16, $p < .05$	β = .19, $p < .01$	β = .12, $p < .05$	β = .31, $p < .01$
	Within-level		β = .13, $p < .05$			
High levels of cooperative climate	Between-level	β = .77, $p < .001$	β = .47, $p < .001$	β = .51, $p < .001$	β = .36, $p < .001$	β = .87, $p < .001$
	Within-level		β = .37, $p < .01$			

Note: HCHRM = high-commitment HRM; high levels of cooperative climate = 1 SD above the mean; low levels of cooperative climate = 1 SD below the mean.

modeling over classical analytical methods. It distills the literature on statistical logic underpinning multilevel modeling, variable level conceptualization, and data aggregation and analysis. An empirical HRM research example is offered to demonstrate the procedures for meeting preconditions for the implementation of multilevel analytical strategies and hypothesis test. A brief assessment of statistical packages for performing multilevel analysis is also provided.

As HLM software is an invaluable, user-friendly computer program for novice users, the procedures for the HLM program setup and hypothesis tests when conducting simple multilevel analysis without cross-level mediation are described in the appendix. In summary, this article provides thorough, actionable knowledge aimed at promoting, and assisting HRM researchers in, conducting multilevel HRM research.

Notes

1. Mathieu et al. (2012) created a computer program that allows researchers to estimate the power of cross-level interactions to determine the sample size of multilevel variables prior to data collection. This program is available online at: http://mypage.iu.edu/haguinis/~-crossless.html.
2. The software is available at http://www.quantpsy.org/selig_preacher_mplus_syntax.htm.
3. Available at http://www.quantpsy.org/interact/hlm2.htm.
4. Such as from http://www.people.ku.edu/~preacher/interact/shacham/index.hlm.

JIE SHEN is associate professor of HRM at the University of South Australia. Previously, he held the positions of full professor of HRM at Curtin University and associate professor of HRM at Monash University. He is an editorial board member of three international journals and the guest editor of the special issue "Multilevel HRM Research" for the *International Journal of HRM.* Since 2004, he has published 7 research books and 55 refereed journal articles, including articles in the *Journal of Management* and *Human Resource Management.*

References

Aiken, L. S., & West, S. G. (1991). Multiple regression: Testing and interpreting interactions. Newbury Park, CA: Sage.

Arnold, C. I. (1992). An introduction to hierarchical linear models. Measurement and Evaluation in Counselling and Development, 25, 58–90.

Aryee, S., Walumbwa, F. O., Seidu, E. Y. M., & Otaye, L. E. (2012). Impact of high performance work systems on individual and branch-level performance: Test of a multilevel model of intermediate linkages. Journal of Applied Psychology, 97, 287–300.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. Journal of Personality and Social Psychology, 51, 1173–1182.

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

Bliese, P. D. (2002). Multilevel random coefficient modeling in organizational research. In F. Drasgow & N. Schmitt (Eds.), Measuring and analyzing behavior in organizations (pp. 401–445). San Francisco, CA: Jossey-Bass.

Bliese, P. D., & Hanges, P. J. (2004). Being both too liberal and too conservative: The perils of treating grouped data as though they were independent. Organizational Research Methods, 7, 400–417.

Bock, G. W., & Kim, Y. G. (2002). Breaking the myths of rewards: An exploratory study of attitudes about knowledge sharing. Information Resources Management Journal, 15(2), 14–21.

Castro, S. L. (2002). Data analytic methods for the analysis of multilevel questions: A comparison of intraclass correlation coefficients, rwg(j), hierarchical linear modeling, within- and between-analysis, and random group resampling. Leadership Quarterly, 13, 69–93.

Chatman, J. A., & Flynn, F. J. (2001). The influence of demographic heterogeneity on the mergence and consequence of cooperative norms in work teams. Academy of Management Journal, 44(5), 956–974.

Collins, C. J., & Smith, K. G. (2006). Knowledge exchange and combination: The role of human resource practices in the performance of high-technology firms. Academy of Management Journal, 49, 544–560.

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

Edwards, J. R., & Lambert, L. S. (2007). Methods for integrating moderation and mediation: A general analytical framework using moderated path analysis. Psychological Methods, 12, 1–22.

Eisenberger, R., Huntington, R., Hutchison, S., & Sowa, D. (1986). Perceived organizational support. Journal of Applied Psychology, 71(3), 500–507.

Freedman, L. S., & Schatzkin, A. (1992). Sample size for studying intermediate endpoints within intervention

trials of observational studies. American Journal of Epidemiology, 136, 1148–1159.
Glick, W. H. (1985). Conceptualizing and measuring organizational and psychological climate: Pitfalls in multilevel research. Academy of Management Review, 10(2), 601–616.
Hofmann, D., Griffin, M., & Gavin, M. (2000). The application of hierarchical linear modeling to organizational research. In K. Klein & S. Kozlowski (Eds.), Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions (pp. 467–511). San Francisco, CA: Jossey-Bass.
Hofmann, D. A. (1997). An overview of the logic and rationale of hierarchical linear models. Journal of Management, 23, 723–744.
Hofmann, D. A., & Gavin, M. B. (1998). Centering decisions in hierarchical linear models: Implications for research in organizations. Journal of Management, 24(5), 623–641.
James, L. R., Demaree, R. G., & Wolf, G. (1984). Estimating within-group inter-rater reliability with and without response bias. Journal of Applied Psychology, 69, 86–98.
Jensen J. M., Patel, P. C., & Messersmith, J. G. (2013). Higher-performance work systems and job control: Consequences for anxiety, role overload, and turnover intentions. Journal of Management, 39, 1699–1724.
Kenny, D. A., Kashy, D. A., & Bolger, N. (1998). Data analysis in social psychology. In D. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), Handbook of social psychology (4th ed., Vol. 1, pp. 233–265). New York, NY: McGraw-Hill.
Kreft, I. G. G., & de Leeuw, J. (1998). Introducing multilevel modeling. Newbury Park, CA: Sage.
Kulik, C. (2012). Climbing the higher mountain: The challenges of multilevel, multisource, and longitudinal research designs. Management and Organization Review, 7(3), 447–460.
LeBreton, J. M., & Senter, J. L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. Organizational Research Methods, 11, 815–852.
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 perspectives of high-performance work systems and influence processes on service quality. Journal of Applied Psychology, 94, 371–391.
Littell, R. C., Milliken, G. A., Stroup, W. W., & Wolfinger, R. D. (1996). SAS systems for mixed models. Cary, NC: SAS Inst.
Liu, D., Liao, H., & Loi, R. (2012). The dark side of leadership: A three-level investigation of the cascading effect of abusive supervision on employee creativity. Academy of Management Journal, 55(5), 1187–1212.
Longford, N. T. (1986). VARCL—Interactive software for variance component analysis: Applications for survey data. Professional Statistician, 5, 28–32.
Luke, D. A. (2004). Multilevel modeling (3rd ed.). Thousand Oaks, CA: Sage.
Maas, C. J. M., & Hox, J. J. (2006). Sufficient sample sizes for multilevel modelling. Methodology, 1(3), 86–92.
MacKinnon, D. P., Fritz, M. S., Williams, J., & Lockwood, C. M. (2007). Distribution of the product confidence limits for the indirect effect: Program PRODCLIN. Behavior Research Methods, 39, 384–389.
Mathieu, J. E., & Chen, G. (2011). The etiology of the multilevel paradigm in management research. Journal of Management, 37, 610–641.
Mathieu, J. E., Aguinis, H., Culpepper, S. A., & Chen, G. (2012). Understanding and estimating the power to detect cross-level interaction effects in multilevel modeling. Journal of Applied Psychology, 97, 951–966.
Messersmith, J. G., Patel, P. C., & Lepak, D. (2011). Unlocking the black box: Exploring the link between high-performance work systems and performance. Journal of Applied Psychology, 96(6), 1105–1118.
Osborne, J. W. (2000). Advantages of hierarchical linear modeling. Practical Assessment, Research, and Evaluation, 7, 1–3.
Ostroff, C., & Bowen, D. E. (2000). Moving HR to a higher level: HR practices and organizational effectiveness. In K. L. Klein & S. W. J. Kozlowski (Eds.), Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions (pp. 211–266). San Francisco, CA: Jossey-Bass.
Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. Journal of Applied Psychology, 88, 879–903.
Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interaction effects in multiple linear regression, multilevel modeling, and latent curve analysis. Journal of Educational and Behavioral Statistics, 31, 437–448.
Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. Behavior Research Methods, 40, 879–891.
Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. Psychological Methods, 15, 209–233.
Raudenbush, S. W., & Bryk, A. S. (2002). Hierarchical linear models: Applications and data analysis methods (2nd ed.). Newbury Park, CA: Sage.
Raudenbush, S. W., & Liu, X. (2000). Statistical power and optimal design for multisite randomized trials. Psychological Methods, 5, 199–213.
Rousseau, D. M. (1985). Issues of level in organizational research: Multi-level and cross-level perspectives. Research in Organizational Behavior, 7, 1-37.
Rowe, K. J., & Hill, P. W. (1998). Modeling educational effectiveness in classrooms: The use of multi-level structural equations to model students' progress. Educational Research and Evaluation: An International Journal on Theory and Practice, 4(4), 307–347.
Snell, S. A., & Dean, J. W. Jr. (1992). Integrated manufacturing and human resource management: A human capital perspective. Academy of Management Journal, 35, 467–504.
Snijders, T. A. B. (1996). Stochastic actor-oriented models for network change. Journal of Mathematical Sociology, 21, 149–172.

Snijders, T. A. B., & Bosker, R. J. (2012). Multilevel analysis: An introduction to basic and advanced multilevel modelling (2nd ed). Los Angeles, CA: Sage.

Sun, L. Y., Aryee, S., & Law, K. S. (2007). Higher-performance human resource practices, citizenship behavior, and organizational performance: A relational perspective. Academy of Management Journal, 50, 558–577.

Takeuchi, R., Chen, G., & Lepak, D. P. (2009). Through the looking glass of a social system: Cross-level effects of high-performance work systems on employees' attitudes. Personnel Psychology, 62, 1–29.

Tse, H. M., Dasborough, M., & Ashkanasy, N. M. (2008). A multilevel analysis of team climate and interpersonal exchange relationships at work. Leadership Quarterly, 19, 195–211.

Wei, Y. C., Han, T. S., & Hsu, I. C. (2010). High-performance HR practices and OCB: A cross-level investigation of a causal path. International Journal of Human Resource Management, 21(10), 1631–1648.

Whitener, E. M. (2001). Do "high commitment" human resource practices affect employee commitment? A cross-level analysis using hierarchical liner modeling. Journal of Management, 27, 515–535.

Willms, J. D., & Raudenbush, S. W. (1989). A longitudinal hierarchical linear model for estimating school effects and their stability. Journal of Educational Measurement, 26, 209–232.

Zheng, Z., Zyphur, M. J., & Preacher, K. J. (2009). Testing multilevel mediation using hierarchical linear models: Problems and solutions. Organizational Research Methods, 12(4), 695–719.

APPENDIX The Procedure for Conducting Multilevel Analysis Using HLM Software

HLM Set-Up

After launching the HLM program, click *File-Make new MDM file* (if using an existing MDM file, click *Create a new model using an existing MDM file*) -*Stat package input-HLM2* (for three level models, click *HLM3*). A dialogue box will open, in which researchers are required to:

1. Create a MDM file using the .*mdm* suffix and save by clicking *Save mdm file*.
2. Specify nesting of input data by choosing *persons within groups*.
3. Load Level-1 and Level-2 files and choose variables. Firstly, choose *Organization* as ID in both files as it links two levels and secondly, choose *Anything else* if using HLM version 6 and *SPSS/Windows* if using HLM version 7.
4. Provide missing data information by choosing *Yes* or *No*. If there is missing data at Level-1, researchers need to specify whether they wish to delete missing data when creating a *mdm* file or running analyses. Missing data are not allowed at Level-2 and Level-3.

It is now possible to create a *mdm* file. Researchers should click *Make MDM* and check the statistics including variables, *n*, mean and SD by clicking *Check Stats*. If satisfied with the statistics, click *Done*. After this process is complete, researchers can perform the required analyses.

Hypothesis Testing

To run the *null model*, click *Level 1* and enter Level-1 dependent variable as *Outcome variable* and click *Run analysis.* To check the results, go to *File* and click *View Output*. To test the intercepts-as-outcomes model, click *Level 1*, enter Level-1 outcome variable as *Outcome variable,* and enter control variables as *uncentered* and Level-1 independent variable as a *group centered*. Then, click *Level 2,* enter Level-2 control variables and the Level-2 independent variable as a *grand centered*.

To test the slopes-as-outcomes model, click the error term for Level-1 variable, after the error term appears, click *Level-2* variable. This will enter an interaction of Level-1 and Level-2 predictors. After all variables are entered, click *Run analysis.* To check the results, go to *File* and click *View Output*.