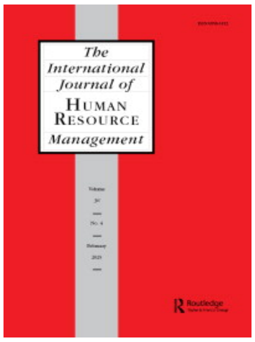

The International Journal of Human Resource Management

ISSN: 0958-5192 (Print) 1466-4399 (Online) Journal homepage: www.tandfonline.com/journals/rijh20

Human resource management systems research – how to gain impactful insights through formative measurement and hierarchical component models

Sven Hauff, Nicole F. Richter & Christian M. Ringle

To cite this article: Sven Hauff, Nicole F. Richter & Christian M. Ringle (2025) Human resource management systems research – how to gain impactful insights through formative measurement and hierarchical component models, The International Journal of Human Resource Management, 36:4, 611-635, DOI: 10.1080/09585192.2025.2464668

To link to this article: https://doi.org/10.1080/09585192.2025.2464668

Published online: 24 Feb 2025.

THE INTERNATIONAL JOURNAL OF HUMAN RESOURCE MANAGEMENT
2025, VOL. 36, NO. 4, 611–635
https://doi.org/10.1080/09585192.2025.2464668

Human resource management systems research – how to gain impactful insights through formative measurement and hierarchical component models

Sven Hauff[a], Nicole F. Richter[b] and Christian M. Ringle[c,d]

[a]Department of Humanities and Social Sciences, Helmut Schmidt University, Hamburg, Germany; [b]Department of Business and Management, University of Southern Denmark, Odense, Denmark; [c]Department of Management Sciences and Technology, Hamburg University of Technology, Hamburg, Germany; [d]College of Business, Law and Governance, James Cook University, Townsville, Australia

ABSTRACT
This paper emphasizes that paying greater attention to how human resource management (HRM) systems are conceptualized in empirical studies could provide more actionable insights and increase the impact of HRM systems research. We advocate the use of formative measurement models, arguing that this approach aligns better with the concept of HRM systems, and allows for a nuanced understanding of how each HRM practice and the system contribute to the outcomes of interest. In the same vein, we advocate the use of hierarchical component models, which allow a multi-level conceptualization representing HRM practices, the HRM system, and their intermediate levels of abstraction (e.g. ability, motivation, and opportunities as subcomponents of high-performance work systems). As a result, HRM systems research can move beyond general assertions and instead offer specific and actionable recommendations. We discuss and illustrate how these conceptual ideas can be implemented in partial least squares-structural equation modeling (PLS-SEM), and enriched by predictive model evaluation following state-of-the-art guidelines.

KEYWORDS
HRM systems; formative measurement; hierarchical component models; predictive model assessment; partial least squares-structural equation modeling; PLS-SEM

Introduction

Research on human resource management (HRM) systems and their associations with work outcomes form the core of strategic HRM (SHRM) (Lepak et al., 2006; Jackson et al., 2014; Jiang & Messersmith, 2018). In spite of the research knowledge generated in the field, several authors have drawn attention to fundamental problems in this research domain

CONTACT Sven Hauff hauff@hsu-hh.de Helmut Schmidt University, Holstenhofweg 85, 22043, Hamburg, Germany.

(Beijer et al., 2021; Boon et al., 2019; Cooke et al., 2021; Hauff, 2021; Kaufman, 2015; Kaufman, 2020). These relate to a missing link to economic outcomes, the omission of the internal and external environmental context, the lack of longitudinal and multilevel studies, and to conceptualization and measurement issues around HRM systems. Among others, these issues contribute to a more general criticism of strategic HRM research for its lack of practice-relevant insights. For example, Cooke et al. (2021, p. 89) observe an 'increasing detachment from HRM practice,' while Kaufman (2020, p. 65) even states that 30 years of HRM systems research has led to 'small-to-negligible knowledge/practice value-added'.

Arguably, the lack of insights relevant to practice is primarily due to the measurement and representation of HRM systems in empirical studies. Boon et al. (2019) report that 69% of the HRM systems they reviewed were measured by using an additive index, an average, or a factor score of HRM practices. These approaches allow the testing of HRM systems' overall effects and have helped to establish an understanding of the general relationships between HRM systems and multiple outcomes (see the meta-review by Jiang & Messersmith, 2018). However, they do not provide meaningful insights into which HRM practices within HRM systems are most important for achieving an outcome. In fact, they only show that 'more HRM is better' (Kaufman, 2020, p. 58) without providing HRM practitioners with support on where to focus their investments.

We argue that SHRM research could provide more impactful insights by paying more attention to how HRM systems are empirically operationalized. More specifically, building on previous research (see Hauff, 2021), we argue that HRM researchers and practitioners could gain more actionable knowledge by using latent constructs with formative measurement models and hierarchical component models (HCMs). Latent constructs with formative measurement models are aligned with the conceptual idea that HRM systems are formative by nature (Jiang et al., 2012a; Hauff, 2021). Importantly, they make it possible to determine how much different HRM practices (or HRM domains) contribute to the overall effect of the HRM system. They provide insights into the influence of the HRM system as a whole as well as the relative influence of the single HRM practices (or HRM domains). Thus, this approach shows the overall effect of the HRM system, and helps to identify the drivers within an HRM system that are influential for a particular outcome. The latter is very important, given that research on the influence of subdimensions of HRM systems has led to concerns that 'the variance explained by HRM systems in general is being driven by a relatively small number of […] practices, rather than the full system' (Ogbonnaya & Messersmith, 2019, p. 520).

HCM allow the integration of constructs measured at a higher level of abstraction, such as the HRM system, while simultaneously including its subcomponents, such as different HRM domains of HRM practices. HCM can therefore help measure HRM systems adequately if these systems are conceptualized at different levels—such as high-performance work systems (HPWS) (Appelbaum et al., 2000; Jiang et al., 2012a) or well-being-oriented HRM systems (Guest, 2017; Gubernator et al., 2024). Also HCM can be of use if a HRM practice is measured by multiple items (e.g. Garmendia et al., 2021).

The SHRM community does not seem to be well acquainted with these methodological opportunities. To identify the prevalence of formative measurement models and HCMs in SHRM research, we reanalyzed the studies presented in the review by Boon et al. (2019) who did not provide any information on the use of formative measurement models and HCM. We found that only 62 (i.e. 12.5%) studies of the 495 empirical studies used a structural equation model (SEM) with latent constructs. Of these studies, 46% used an HCM, while only 6.5% used a formative measurement model. In total, only 0.8% of all the studies included in their review used a latent construct approach with a formative measurement model to operationalize HRM systems.

Against this background, our aim is to bridge this knowledge gap and promote a greater adoption of formative measurement models and HCMs in SHRM research. We support this by discussing in detail how formative measurement models and HCMs fit with key ideas in HRM systems research. We also illustrate how to realize these ideas in partial least squares structural equation modeling (PLS-SEM) (Wold, 1982; Lohmöller, 1989), and provide an overview of state-of-the art guidelines for their evaluation and the assessment of explanatory and predictive power. Ultimately, this may help to improve the conceptual accuracy, methodological rigor and practical relevance of HRM systems research.

HRM systems and their measurement

HRM systems represent ‘bundles of HR practices intended to achieve the objectives of organizations’ (Jiang et al., 2012a, p. 73). Because HRM can pursue different goals, SHRM research refers to several variations of HRM systems designed to achieve specific goals. The most prominent HRM systems are HPWS, high-commitment work systems (HCWS), and high-involvement work systems (HIWS). Less frequently studied HRM systems are relationship-oriented HRM systems, knowledge-oriented HRM systems, or initiative-enhancing HRM systems (Boon et al., 2019).

Researchers mostly use an additive index, an average score, or a factor score of HRM practices to measure HRM systems (Boon et al., 2019).[1]

Additive indices measure the presence of (coverage of, satisfaction with) specific HRM practices (e.g. training and development, performance appraisal, and autonomy), and sum all the responses (usually on Likert scales) to a single index, which is then used in the empirical analysis. To determine an average score, this sum is divided by the number of HRM practices. Researchers typically use some form of factor analysis to create factor scores in order to demonstrate their HRM system measure's reliability and to aggregate individual HRM practices into a single factor score by using principal component analysis. Researchers usually do not provide a justification for the use of a particular approach; they state that they are '[f]ollowing prior research' (Zhou et al., 2023, p. 815), that their aggregation is 'in line with other SHRM studies' (Kim, 2024, p. 2235) or that they use a scale from previous studies (Alothmany et al., 2023).

Additive indices, average scores, and factor scores of HRM practices combine different HRM practices into an overall measure to analyze the HRM system's impact on a variety of outcomes. These overall measures involve, however, several methodological problems. First, by simply summing individual HRM practices, each practice is weighted equally, not taking their relative influence into account (e.g. Hauff, 2021). Additive indices also overlook measurement theory, exhibit lower construct validity and reliability, and may result in inaccurate coefficient estimates, potentially leading to erroneous conclusions (McNeish & Wolf, 2020; McNeish, 2023; Hair, Sharma, Sarstedt, et al., 2024). Second, factor analyses' application could be criticized for not suiting the HRM system's formative nature, which several authors highlighted (e.g. Jiang et al., 2012a; Jiang & Messersmith, 2018) (see next paragraph). Third, all these approaches only capture the overall relationship between the HRM system and the outcomes of interest, showing whether 'more HRM is better,' but do not provide practical recommendations regarding specific HRM domains or practices.

These issues can be addressed by using latent constructs with formative measurement models. Indicators in reflective measurement models are manifestations of the construct, meaning that all indicators in the model are intended to reflect the same construct. Consequently, the indicators are typically interchangeable and expected to be highly correlated. A formative measurement model aims to cover all aspects that cause, contribute to, or form the latent construct (i.e. the HRM system). Each formative indicator (i.e. the individual HRM practices) is assumed to capture a salient aspect of the latent construct. Consequently, formative indicators are not interchangeable, and there are no specific expectations of their correlations; they could even be completely independent (MacKenzie et al., 2005; Diamantopoulos et al., 2008; Sarstedt et al., 2016). The formative approach is consistent with the idea that HRM practices serve different purposes, and are not necessarily used in conjunction with one another. For example, training activities aim to increase (i.e. contribute to) employees' knowledge and skills, while incentive-based pay aims to increase their motivation (Bos-Nehles et al., 2023). While organizations may pursue both goals, there is no guarantee that training and pay-for-performance are always used together (and are therefore highly correlated), since some organizations may rely on intrinsic motivation instead of pay-for-performance. Notably, if each indicator is assumed to measure a salient aspect of the latent construct, individual HRM practices should only appear once in the formative measurement model. Consequently, if multiple items are used to measure a single HRM practice (e.g. multiple items to measure training), they should either be combined into a single item (e.g. by averaging the responses from the items) or the researcher should consider an HCM (see below).

Formative measurement models' nature allows researchers to understand each indicator's individual contributions to both the construct and its outcomes by estimating the formative indicator weights. These represent how much each HRM practice in an HRM system contributes to the overall effect. This information empowers HRM systems research by offering insights into the HRM practices that, in addition to the HRM system's influence, are most critical in respect of achieving the desired outcomes. Thus, researchers who want to go beyond simply understanding whether more HRM is better for an outcome can benefit from formative measurement models, as these models allow for an understanding of which HRM practices specifically influence the outcome.

Importantly, formative measurement models require researchers to ensure that their indicators capture all (major) facets of their construct's domain. Based on theoretical considerations, these indicators should include a comprehensive set of HRM practices that are salient and influential for the goal of the HRM system. For example, if the HRM system's aim is to increase employee performance (as is the aim of HPWS), researchers need to include all (major) HRM practices that are supposed to increase employee performance (e.g. based on the ability, motivation, and opportunity framework). The critical importance of establishing a clear link between various HRM theories, related HRM objectives, and corresponding sets of HRM practices has been recently emphasized by Guest (2025). He advocates 'a more careful, theoretically informed measurement of HR practices by aligning the chosen practices with the underlying theory being tested' (ibid., p. 12). Formative measurement models should trigger this careful measurement model operationalization as they require that researchers define the specific type of HRM system being studied, along with its intended goals, in advance. Only this conceptual clarity enables an appropriate consideration of all relevant HRM

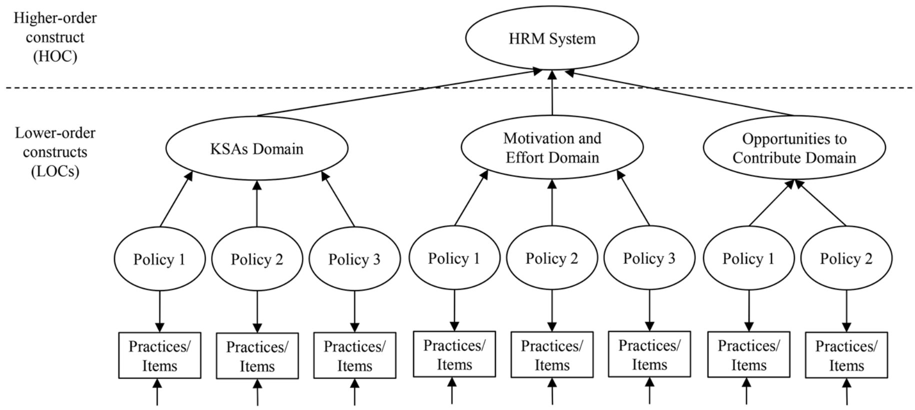

Figure 1. Higher and lower order components in HRM systems.
Note: KSAs = knowledge, skills, and abilities.
Source: Adapted from Jiang, Lepak, Han, et al. (2012a).

practices. This process also helps to avoid unspecific labels (e.g. 'HRM system' instead of HPWS), which can lead to ambiguous measures of low value (Boon et al., 2019). This also implies that there is not a single 'ideal' measure of all HRM systems since the measurement must always be tailored to the specific goal and the context of the HRM system.

HRM systems can be conceptualized as first-order models with the HRM system being directly measured *via* indicators on a single level. Alternatively, they can be conceptualized in a more complex structure that includes different levels of abstraction (also referred to as hierarchical component models or HCMs). In such a structure, the higher-order constructs (HOCs) represent the highest level of abstraction, and the lower-order components (LOCs) represent more concrete subdimensions. Jiang et al. (2012a) have used such a structure to conceptually describe the measurement of HRM systems (see Figure 1). In these scholars' approach, the HRM system is the highest level of abstraction (i.e. the HOC). Building on the ability, motivation, and opportunity (AMO) framework (e.g. Appelbaum et al., 2000; Lepak et al., 2006; Bos-Nehles et al., 2023), three domains form the HRM system (i.e. knowledge, skills, and abilities, motivation and effort, and opportunities to contribute), which in turn are formed by even more concrete HRM policies.[2]

The relationships between the HOCs and the LOCs can conceptually be either reflective or formative, leading to four types of HCMs: reflective-reflective, reflective-formative, formative-reflective, and formative-formative HCMs (Sarstedt et al., 2019; Becker et al., 2023; Hair, Sarstedt, Ringle, et al., 2024). A formative-formative model is appropriate when HRM systems are formed by different domains which are in turn formed by HRM practices. A reflective-formative measurement is appropriate when researchers want to measure the HRM systems with different HRM practices, which are, in turn,

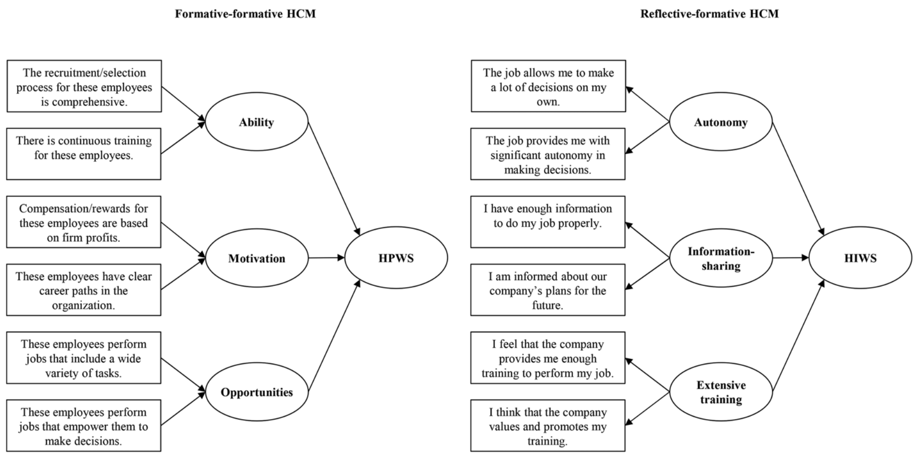

Figure 2. Formative-formative vs. formative-reflective HCM.
Note: Ovals represent constructs that are not directly measured. Rectangles represent the indicators as directly measured variables. Arrows pointing from the indicators to the construct indicate formative relationships or measurements, while arrows pointing from the construct to the indicators indicate reflective relationships or measurements; HPWS = high-performance work system; HIWS = High-involvement work system; HCM = hierarchical component model.

measured with multiple items. Given that HRM systems are formative by nature (see the discussion above), the reflective-reflective and the formative-reflective measures do not seem appropriate for HRM systems research, as a reflective measurement mode at the highest level of the HOC does not fit the idea of a formative conceptualization.

Figure 2 demonstrates the distinction between a formative-formative and a reflective-formative HCM. The formative-formative HCM shows an excerpt of the HPWS measure, which we use later in our empirical example. Following Jiang et al. (2012a), the HPWS is based on the AMO framework (e.g. Appelbaum et al., 2000; Lepak et al., 2006; Bos-Nehles et al., 2023). We therefore distinguish between ability, motivation, and opportunity as separate domains that should enhance employee performance. These domains measure the HPWS formatively, because ability, motivation, and opportunity are distinct drivers of employee performance. All three domains are formatively measured by the corresponding HRM practices; that is HRM practices that should enhance employees' abilities, motivation and opportunities respectively (Jiang, Lepak, Hu, et al., 2012b). The reflective-formative HCM illustrates selected measures for HIWS proposed by Garmendia et al. (2021). Here, the HRM practices (i.e. autonomy, information sharing, extensive training) measure the HRM system formatively. Each practice is measured reflectively with items that are (interchangeable) manifestations of these practices (e.g. see the indicators measuring autonomy). Using multiple reflective (instead of single) indicators could increase the measurement accuracy

(Diamantopoulos et al., 2012). Ultimately, theory should always guide the question of how to conceptualize the HCM.

PLS-SEM capabilities regarding HRM systems' research

The PLS-SEM method

PLS-SEM has been widely adopted in business and management research (e.g. Petter & Hadavi, 2023) both as an alternative to covariance-based structural equation modeling (CB-SEM, Jöreskog, 1982; Sarstedt et al., 2016) and a complementary method to CB-SEM (Jöreskog & Wold, 1982; Rigdon et al., 2017; Sarstedt et al., 2024). When using PLS-SEM, researchers establish a structural model that specifies the causal-predictive relationships between constructs (i.e. their theory's unobserved and not directly measurable elements; e.g. HPWS) (Chin et al., 2020; Sarstedt et al., 2021). To estimate the dependent (endogenous) and independent (exogenous) constructs, researchers use indicators or observed variables and theoretically distinguish between reflective and formative measurement models (Sarstedt et al., 2016; Bollen & Diamantopoulos, 2017).

PLS-SEM allows the easy incorporation of formative measurement models and HCM (Sarstedt et al., 2019; Becker et al., 2023; Legate, Ringle, et al., 2023). To help SHRM researchers make more use of formative measures and HCM, we summarize the guidelines for implementing formative and HCM in PLS-SEM, and also provide insights into the latest developments regarding predictive model assessment, which could benefit SHRM research. We hope that this will also help avoid the misapplications found in other fields (Ringle et al., 2020; Sarstedt et al., 2022).

Modelling of formative measurement constructs

While CB-SEM can handle formative indicators, it requires adherence to specific specification rules to ensure model identification (Fornell & Bookstein, 1982; Cenfetelli & Bassellier, 2009), which often leads to challenges when executing the analysis as initially intended (Diamantopoulos & Riefler, 2011). Conversely, PLS-SEM accommodates formative measurement models with less restrictive constraints. Formative measurement models should, however, not be evaluated by using reflective measurement models' standard evaluation criteria (e.g. Cronbach's alpha, MacKenzie et al., 2005). Instead, the evaluation relies on an assessment of the collinearity of indicators, and the statistical significance and relevance of their weights (Hair et al., 2019).[3]

To evaluate the formative indicators' *collinearity*, we check whether the variance inflation factors (VIF) of each indicator are close to 3 or lower (Hair et al., 2019). Next, we evaluate whether each indicator contributes to the index by assessing the indicator weights' *significance and relevance*. The values of the outer weights indicate each indicator's relative contribution to the construct. We assess whether formative indicators truly contribute to forming the construct by testing if the outer weights are significantly different from zero. The significance testing builds on bootstrap confidence intervals and the percentile method (if the bootstrap distribution of the outer weights is highly skewed, we use the bias-corrected and accelerated (BCa) bootstrap confidence intervals) (Aguirre-Urreta et al., 2018). Insignificant indicator weights do not necessarily imply poor measurement model quality. We also need to consider a formative indicator's absolute contribution to its construct (i.e. its loading or its bivariate correlation with its construct) (Cenfetelli & Bassellier, 2009). Only when both the weights and the loadings are insignificant, and the absolute contribution is low (i.e. below 0.5), researchers should consider deleting a formative indicator from the measurement model. Researchers need to be very cautious when removing indicators from the construct, since, as defined in the conceptualization stage, formative measurement theory requires the indicators to fully capture a construct's domain (Hair et al., 2022).

Modeling of higher-order constructs

Another PLS-SEM strength is its ability to handle complex models, including HCM. To identify a HCM in PLS-SEM, researchers can apply repeated indicator or two-stage approaches (Hair et al., 2022). While these approaches' results and those of their sub-forms are often highly similar, the latest research recommends using the disjoint two-stage approach. This approach offers advantages regarding specific model constellations (compared to the repeated indicators approach), allowing researchers to employ all prediction-related PLS procedures (compared to the embedded two-stage approach) (Becker et al., 2023). This flexibility is particular relevant in HRM research, which calls for more predictive power assessments (Sarstedt & Danks, 2022).

In the first stage of the disjoint two-stage approach, construct scores are built that can be used to measure the HOCs in the subsequent stage. We do so by using the LOCs and connecting them to the HOCs' antecedents and consequences in the model (i.e. to the determinants, if included in the model, and outcomes of the HRM system). The standard PLS-SEM algorithm provides construct score estimates for the LOCs based on the modeling described above. These LOC construct scores are used in the

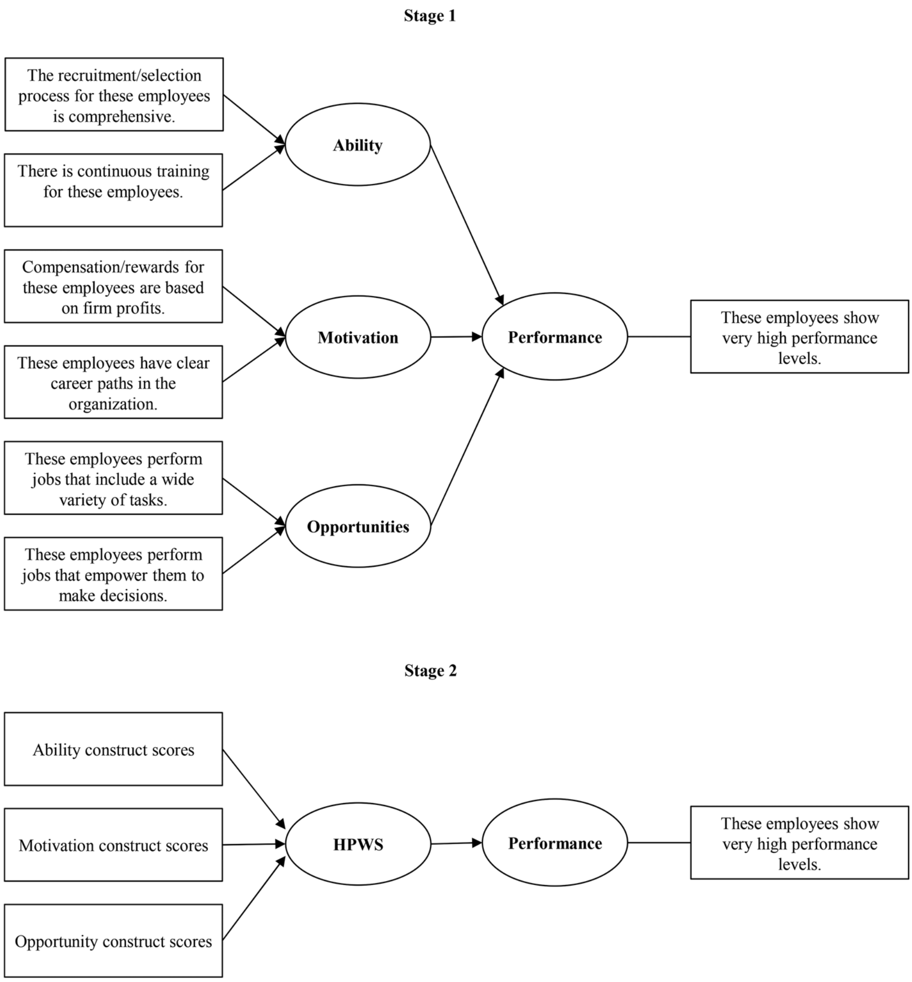

Figure 3. Stages of the disjoint two-stage approach.

second stage to measure the HOC, while all the other (nonhierarchical) constructs are measured by using their original indicators (Becker et al., 2023). Figure 3 illustrates the modeling of the disjoint two-stage approach's first and second stages.

In an HCM, it is imperative to evaluate the measurement model's quality in both stages. This entails ensuring adequate quality of the measurement model for LOCs along the standard evaluation criteria for formative and reflective measurement models, respectively, before proceeding to the second stage. In the second stage, all construct measures, including the HOC, must be evaluated (Hair et al., 2019; Becker et al., 2023). The relationships between the HOC and LOC, regardless of whether they are formative-formative, or reflective-formative HCM, are interpreted as weights. Consequently, the assessment of such HCMs follows the evaluation of formative measurement models.

Assessing model quality and its predictive performance

Explanatory power refers to a model's ability to test the underlying mechanisms and causal relationships between inputs and outcomes (Shmueli, 2010). In this context, researchers often refer to the coefficient of determination (R^2); that is the amount of explained variance in the outcome construct (e.g. how much of the variance in employee performance is explained by the HRM system) (Hair et al., 2022).

Predictive power refers to the ability of a model or a theory to accurately forecast (future) outcomes based on available (past) data (Sharma et al., 2024). For instance, researchers forecast the employee performance resulting from different HRM systems. Researchers who solely focus on obtaining higher R^2-values to optimize their models, may overfit the model to the sample data, thereby, potentially reducing its out-of-sample predictive power (Liengaard et al., 2021). The latter is particularly important for assessing the model's practical relevance. It ensures that recommendations derived from the estimated model address real-world HRM challenges effectively—a neglected, but crucial, assessment in HRM research (Sarstedt & Danks, 2022). To assess predictive power, researchers should use the $PLS_{predict}$ routine (Shmueli et al., 2016), which uses a cross-validation procedure. It estimates the model parameters (e.g. path coefficients, indicator weights) using a training sample, and afterwards applies the model to predict the indicator scores of the outcome construct(s). The predicted values for cases in a holdout or validation sample are subsequently compared to the indicators' actual values in the raw data (Shmueli et al., 2019), and along metrics such as the root mean square error (RMSE), and the mean absolute error (MAE) (for an overview, see Richter & Tudoran, 2024).

PLS-SEM research has developed benchmarking routines, test statistics, and decision trees (e.g. Shmueli et al., 2019) to evaluate the predictive power of PLS-SEM models: In order to demonstrate a high predictive power, the PLS-SEM results should show a lower prediction error compared to the prediction based on the indicator averages (IA; i.e. in our case, the HRM practices' average scores) and a lower prediction error compared to the predictions *via* a linear model (LM) with all indicators (i.e. in our case, a regression of all the individual HRM practices on employee performance). Formally, it should have a $Q^2_{predict}$ value above 0 (indicating that it outperforms the IA benchmark) and an RMSE or MAE value lower than that of the LM. If PLS-SEM only beats the IA benchmark, the model at best has a weak predictive power. If it beats none of the prediction benchmarks, it does not have a suitable predictive power. These findings can be supported by the cross-validated predictive ability test (CVPAT) (Liengaard et al., 2021; Sharma et al., 2023). This test allows assessing and comparing models' predictive power (either that of benchmark models or of alternative conceptual models): It refers to the loss difference (or prediction error) between the models when predicting the indicators of all (overall model), or of a specific endogenous construct (or constructs). The model with the smaller loss (the better prediction error) is preferred. If the loss of Model 1 minus the loss of Model $2 \leq 0$, then Model 1 is preferred; if the loss of Model 1 minus the loss of Model $2 > 0$ and is statistically significant (e.g. with $p < 0.05$), then Model 2 is preferred.

An optional step in predictive model assessment is the selection of competing theoretical models. In order to select a suitable model, authors propose first comparing the alternative models' Bayesian information criterion (BIC), which optimizes the predictive accuracy, and balances the model fit and the complexity to avoid overfitting (Sharma et al., 2019). The comparison along the BIC differences can be complemented by BIC-based Akaike weights which provide insights on a model's relative likelihood (Danks et al., 2020; Rigdon et al., 2023). The model with the highest relative likelihood fits the data best. Finally, researchers can also compare alternative models along the CVPAT statistic as described above.

Table 1 summarizes the assessment criteria of HRM systems with an HCM. The assessment criteria of formative measurement models and the structural model also apply to HRM systems without an HCM.

Illustrative application of PLS-SEM in HRM systems research

Data

For our illustrative application, we utilized sample data sourced from Hauff (2021). It comprises responses from 1,099 chief executives and HR managers of German firms with a minimum of 20 employees from four different sectors (chemicals and pharmaceuticals, mechanical engineering, banking and insurance, and professional services). The respondents provided insights into the employee group deemed most critical for the firm's economic success. For further details, see Hauff (2021).

Measures

We measured the dependent construct with a single *employee performance* indicator ('These employees show very high-performance levels') and a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Given our emphasis on a broad employee performance criterion, this single item enables a reliable and valid assessment of the overall situation by the respondents (see also, e.g. Diamantopoulos et al., 2012). Moreover, it has the advantage of being applicable to different organizational contexts and aligns well with the responsibilities of chief

Table 1. PLS-SEM assessment criteria for an HRM system with an HCM.

Stage	Formative-formative HRM system	Reflective-formative HRM system
Stage 1: LOCs' measurement models	*Formative measurement model* • Collinearity between indicators: The variance inflation factor (VIF) < 3. • Significance and relevance of formative indicators: Larger and significant weights contribute more; maintain indicators with non-significant weights if their indicator loading ≥ 0.5 and significant.	*Reflective measurement model* • Indicator reliability: Outer loadings ≥ 0.708. • Internal consistency reliability: • Cronbach's alpha, reliability ρ_A, and the composite reliability ρ_C 0.70–0.90. • Convergent validity: The average variance extracted (AVE) > 0.50. • Discriminant validity: Heterotrait-monotrait (HTMT) ratio < 0.85–0.90 incl. test if significantly lower than the threshold.
Stage 2: The HOC's measurement model; other constructs' measurement models (see criteria from Stage 1): Structural model assessment	*Formative measurement model (of the HOC)* • Collinearity between the indicators (here the LOC's construct scores): VIF < 3. • Significance and relevance of formative indicators (here the LOC's construct scores): Larger and significant weights contribute more; maintain indicators with non-significant weights (here, the LOC's construct scores) if indicator loading ≥ 0.5 and significant *Structural model* • Collinearity: The VIF between the structural model constructs < 3. • Path coefficients' significance and relevance in the structural model. • Explanatory power: R^2 values. • Predictive power: $Q^2_{predict} > 0$, PLS-SEM root mean squared errors (RMSE) < RMSE of the linear model benchmark for a minority, majority or all indicators of the dependent construct are indicative of low, medium, or high predictive power. • Cross-validated predictive ability test (CVPAT): The model with the smaller loss (a better prediction error) is preferred. If the loss of Model 1 minus the loss of Model 2 ≤ 0, then Model 1 is preferred; if the loss of Model 1 minus the loss of Model 2 > 0 and is statistically significant (e.g. with $p < 0.05$), then Model 2 is preferred.	
Optional additional assessment: Comparing alternative conceptual models	• Bayesian information criterion (BIC): Smaller values indicate higher predictive power; values depend on the scaling of the observed or latent variables used. • Akaike weights: The highest relative likelihood indicates the best fitting model. • CVPAT: See above.	

Source: Based on Hair et al. (2019), Hair et al. (2022), and Sarstedt et al. (2019). *Note*: HOC = higher-order construct, HCM = hierarchical component model, LOC = lower-order construct.

executives and HR managers, who typically evaluate employee performance at the group level.

We refer to *HPWS* as our independent construct representing the most prominent HRM system. It was measured with 10 indicators covering ability-, motivation-, and opportunity-enhancing high-performance work practices (HPWP) as identified by Jiang, Lepak, Hu, et al. (2012b) (see Table 2 for an overview). Respondents used a 5-point Likert scale to indicate their agreement with these statements.

Analysis

We use SmartPLS 4 (Ringle et al., 2024) with the following algorithm settings for our analysis: path weighting scheme, 3,000 iterations,

Table 2. Measures of high-performance work practices.

Domain	HPWP practices	Measures
Ability-enhancing HPWP	Comprehensive recruitment/ selection	The recruitment/selection process for these employees is comprehensive.
	Continuous training	There is continuous training for these employees.
Motivation-enhancing HPWP	Profit-based pay	Compensation/rewards for these employees are based on firm profits.
	Extensive benefits	Compensation/rewards for these employees include an extensive benefits package.
	Clear career paths	These employees have clear career paths in the organization.
	Job security	These employees have long-term perspectives.
Opportunity-enhancing HPWP	Task variety	These employees perform jobs that include a wide variety of tasks.
	Semi-autonomous work groups	These employees work in semi-autonomous work groups.
	Empowerment	These employees perform jobs that empower them to make decisions.
	Information-sharing	Superiors and employees engage in intensive information exchange.

stop-criterion 0.0000001, and mean-replacement for missing values. To obtain bootstrapping results, we used 10,000 bootstrap samples, the no sign change option, percentile bootstrapping, two-tailed testing, and a significance level of 0.05. Building on Hauff (2021), we indicate two models: In Model 1, we use all the high-performance work practices (HPWPs) as indicators of the HWPS; that is, we do not distinguish between LOCs and a HOC. In Model 2, we distinguish between the LOCs and the HOC. We conceptualize the HRM system as a formative-formative HOC with the same ability, motivation, and opportunity domains as the LOCs (see Figure 4). Estimating the higher-order model requires the disjoint two-stage approach (Becker et al., 2023).

Results

First, we evaluate the measurement models. Our single item for employee performance doesn't necessitate a measurement model assessment. Consequently, we focus on evaluating the HRM systems' formative measurement models by following the guidelines discussed above. That is, we check for collinearity issues by examining the VIF values and assessing the formative indicators' significance and relevance.

In Model 1, the indicators' VIF values are below 3, which indicates that there is a low likelihood of multicollinearity issues. Five indicators have insignificant weights (see Table 3). In Model 2, we first evaluate the measures of the LOCs from the first stage and then the measures of the HOC from the second stage of the two-stage approach. The VIF values in the first stage are below 3, with two indicators showing an insignificant weight. The VIF values in the second stage are also below 3, and all the AMO domains have a significant weight. We considered removing

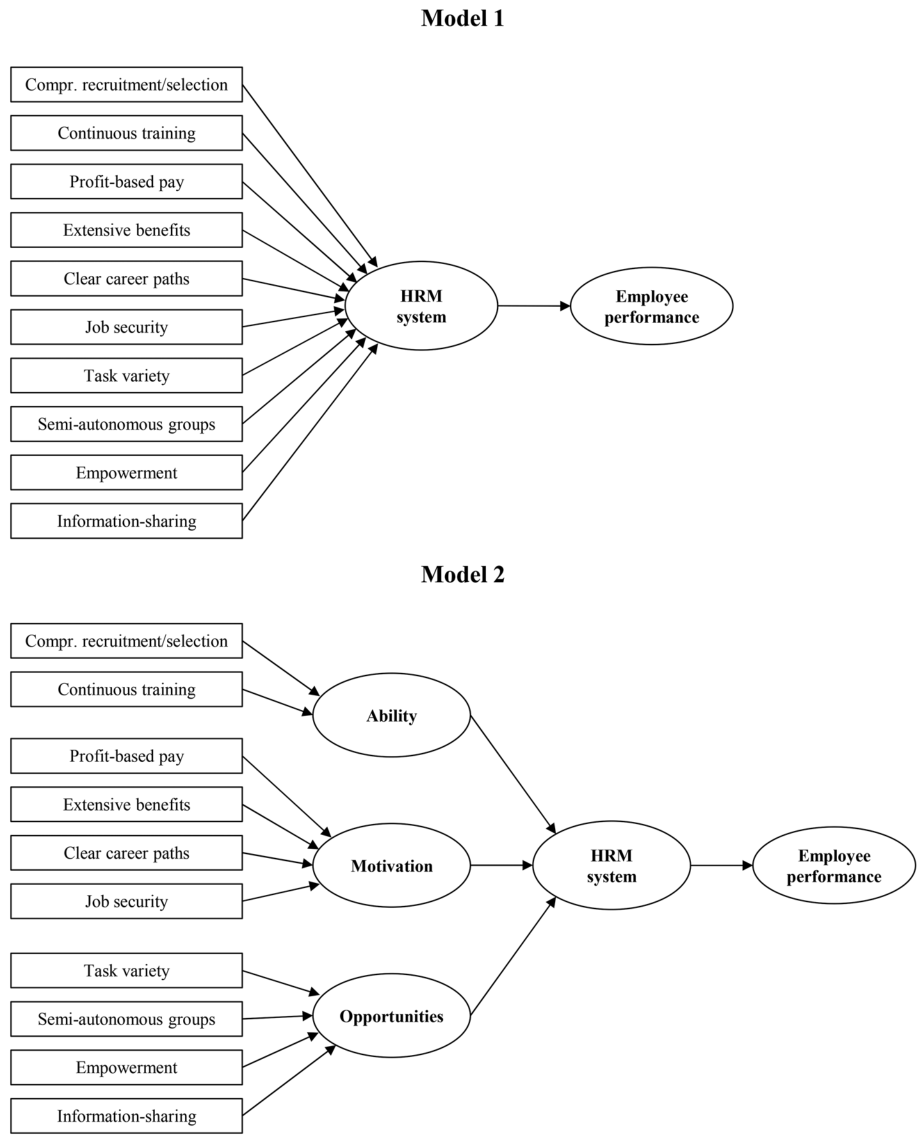

Figure 4. Measurement models.

the indicators with an insignificant weight, specifically if the respective loading is below 0.5. However, given that the respective HPWPs (i.e. continuous training, profit-based pay, extensive benefits, clear career paths, semi-autonomous work groups in Model 1, as well as extensive benefits, and semi-autonomous work groups in Model 2) are theoretically well established regarding their relevance for a HPWS (e.g. Jiang, Lepak, Hu, et al., 2012b; Posthuma et al., 2013), we retain them in the measurement model.

Next, we evaluate the structural model results according to the discussed explanatory and predictive power assessments. Since we only have

Table 3. Indicator weights, loadings, and their significance.

	Model 1				Model 2			
HPWP	Weights (Loadings)	p value	95% Percentile Confidence Interval	Significant?	Weights (Loadings)	p value	95% Percentile Confidence Interval	Signifi-cant?
Ability					0.220 (0.579)	0.002 (0.000)	[0.077, 0.355]	Yes
Comprehensive recruitment/ selection	0.209 (0.462)	0.003 (0.000)	[0.073, 0.348]	Yes	0.636 (0.811)	0.000 (0.000)	[0.413, 0.822]	Yes
Continuous training	0.120 (0.451)	0.092 (0.000)	[−0.019, 0.260]	No	0.610 (0.793)	0.000 (0.000)	[0.381, 0.800]	Yes
Motivation					0.342 (0.667)	0.000 (0.000)	[0.191, 0.492]	Yes
Profit-based pay	0.111 (0.160)	0.094 (0.026)	[−0.012, 0.247]	No	0.258 (0.240)	0.012 (0.027)	[0.063, 0.462]	Yes
Extensive benefits	0.031 (0.201)	0.620 (0.003)	[−0.090, 0.152]	No	0.113 (0.302)	0.260 (0.002)	[−0.088, 0.303]	No
Clear career paths	0.057 (0.365)	0.411 (0.000)	[−0.080, 0.195]	No	0.405 (0.546)	0.000 (0.000)	[0.210, 0.601]	Yes
Job security	0.344 (0.575)	0.000 (0.000)	[0.202, 0.502]	Yes	0.793 (0.861)	0.000 (0.000)	[0.649, 0.911]	Yes
Opportunity					0.712 (0.905)	0.000 (0.000)	[0.576, 0.824]	Yes
Task variety	0.349 (0.650)	0.000 (0.000)	[0.221, 0.485]	Yes	0.472 (0.730)	0.000 (0.000)	[0.335, 0.612]	Yes
Semi-autonomous work groups	0.058 (0.285)	0.393 (0.000)	[−0.073, 0.192]	No	0.105 (0.320)	0.172 (0.000)	[−0.046, 0.255]	No
Empowerment	0.200 (0.591)	0.004 (0.000)	[0.066, 0.339]	Yes	0.319 (0.664)	0.000 (0.000)	[0.173, 0.465]	Yes
Information-sharing	0.359 (0.683)	0.000 (0.000)	[0.238, 0.489]	Yes	0.535 (0.766)	0.000 (0.000)	[0.402, 0.669]	Yes

Note: Individual indicators weights (i.e. HPWP) in Model 2 in the first stage of the two-stage approach, and bias-corrected confidence intervals; HPWP = high-performance work practice.

Table 4. PLS$_{predict}$ results.

	$Q^2_{predict}$	PLS-SEM RMSE	LM RMSE	PLS-SEM MAE	LM MAE
Model 1					
Employee performance	0.178	0.614	0.614	0.479	0.479
Model 2					
Employee performance	0.182	0.612	0.612	0.475	0.475

Note: PLS-SEM = partial least squares structural equation modeling; RMSE = root mean square error; LM = linear model; MAE = mean absolute error.

Table 5. CVPAT results.

		Losses	Average loss differences	*p* values
Model 1		*0.377*		
Indicator average (IA)	Employee performance	0.458	−0.081	0.000
Linear model (LM)	Employee performance	0.377	0.000	0.855
Model 2		*0.375*		
Indicator average (IA)	Employee performance	0.458	−0.083	0.000
Linear model (LM)	Employee performance	0.375	0.000	0.727

one independent construct, collinearity is not an issue in our structural model. In both models, the path coefficient between the HRM system and performance is significant (Model 1: 0.442, $p < 0.001$, 95% CI [0.382, 0.483]; Model 2: 0.433, $p < 0.001$, 95% CI [0.379, 0.481]). To assess the *explanatory power*, we examine the R^2 values of the endogenous construct (i.e. employee performance). The R^2 value is 0.195 in Model 1 and 0.188 in Model 2, which we consider satisfactory (for comparison, R^2 values in Jiang, Lepak, Hu, et al.'s (2012b) meta-analysis ranged between 0.15 and 0.38).

Next, we assess the models' *predictive power* using PLS$_{predict}$ with 10 folds and 10 repetitions (see Shmueli et al., 2019 for more detailed explanations). In both models, the $Q^2_{predict}$ value of the employee performance indicator is larger than zero (0.178 in Model 1, 0.182 in Model 2; Table 4). The models have a significantly higher predictive power than the IA prediction benchmark, as the CVPAT indicates (see Table 5). However, the comparison of the PLS-SEM model and the LM benchmark according to the RMSE and MAE values reveals that the PLS-SEM outcomes do not exhibit statistically significant superiority over the LM benchmark (see Table 5). Consequently, both models individually have a low predictive relevance since the outcomes can only beat the more naïve IA, but not the more demanding LM prediction benchmark.

To compare both models' overall structure with one another, we examine their BIC values and BIC-based Akaike weights. Model 1 has a lower BIC than Model 2, and a clearly higher relative likelihood of 99.5% (Model 1: BIC = −225.765, Akaike weight = 0.995, Model 2: BIC = −215.362, Akaike weight = 0.005), indicating its superiority. We also perform a CVPAT to test if Model 1 has a significantly better predictive power than

Table 6. Total effects.

HPWP	Model 1	Model 2
Comprehensive recruitment/ selection	0.092	0.061
Continuous training	0.053	0.058
Profit-based pay	0.049	0.038
Extensive benefits	0.014	0.017
Clear career paths	0.025	0.060
Job security	0.152	0.117
Task variety	0.154	0.146
Semi-autonomous work groups	0.026	0.032
Empowerment	0.088	0.098
Information-sharing	0.159	0.165

Note: Total effects as the product of the weights and the path coefficients; in Model 2, weights from the first stage were multiplied with the weights of the LOCs and the path coefficients in the second stage; HPWP = high-performance work practice.

Model 2, which is, however, not the case (the loss difference of −0.002 between the two models is not significant; *p*value: 0.438).

In a final step of our analysis, we determine the relative importance of each HRM practice for the outcome in question. This was done by multiplying each HRM practice's weights by the path coefficient(s) of the latent construct(s), which gives each HRM practice's total effect (see Table 6). In both models, information sharing, task variety, and job security are the three most important HPWPs for employee performance.

Discussion and conclusion

In response to recent criticisms of the disconnect between strategic HRM research and its practical applicability (Beijer et al., 2021; Boon et al., 2019; Cooke et al., 2021; Hauff, 2021; Kaufman, 2015; Kaufman, 2020), we argue that the relevance and applicability of HRM systems research can be significantly improved by paying more attention to HRM systems' operationalization in empirical studies.

First, this relates to the use of latent constructs with formative measurement models. These better fit the conceptual foundations of HRM systems (Jiang et al., 2012a; Hauff, 2021), as different HRM practices or domains form these systems. Empirically, they provide insights into how much each HRM practice or domain contributes to the HRM system and its outcomes, moving the field above and beyond general 'more HRM is better' statements toward actionable recommendations for effective HRM practice. We illustrated this by using an HPWS example. Using a formative measurement approach, we found evidence of the relationship between the HPWS and employee performance, and gained insights into the prominent role that information sharing, task variety, and job security play in influencing employee performance. These more actionable insights can help address the challenges that organizations face today.

Second, this relates to the use of HCM. HCM should be used when the theoretical conceptualization of HRM systems comprises multiple levels, which applies to various HRM systems. The AMO model—which is increasingly becoming a standard for measuring an HPWS (Boon et al., 2019)—seems predestined in this respect, but there are also other conceptualizations of HRM systems that refer to multiple dimensions within the overall systems. For example, in his well-being-oriented HRM approach, Guest (2017) describes five domains (i.e. investing in employees, providing engaging work, a positive social and physical environment, voice, and organizational support), each of which includes multiple HRM practices to enhance employee well-being (see also Gubernator et al., 2024). Empirical designs should reflect these conceptual ideas to empirically support and potentially advance these conceptualizations. We illustrated this by using our HPWS example. AMO domains formed the HPWS, but were themselves formed by different HPWP. Our model provided insights into the HPWS's relevance as a whole for employee performance, and into the role that each AMO domain plays. Furthermore, our model identified individual HRM practices' roles.

HCMs are also beneficial when researchers use multiple items rather than single items to measure their HRM practices, which is not uncommon in HRM systems research. For example, Garmendia et al. (2021) used three items to measure training, participation, autonomy, and information, respectively. Whether there is a need to measure HRM practices with multiple items is, however, debatable. On the one hand, it could be argued that HRM practices are rather concrete, which makes multi-item measures (which affect practical considerations regarding the questionnaire's length) superfluous. On the other hand, it could be argued that HRM practices are rather complex phenomena, especially regarding their general principles. Resolving this issue is a task for future research (for general guidelines, see Diamantopoulos et al., 2012). Models that allow testing for the additional value of multi-item measures, such as HCM, are beneficial for advancing the field in this regard.

The utility of practical recommendations is maximized when empirical models have strong predictive performance, which differs from the explanatory power of a model that relates to the data at hand. Models with a high explanatory power do not necessarily exhibit a high predictive accuracy regarding new practical applications. Therefore, researchers call for a consideration of both, and Sarstedt and Danks (2022) have recently emphasized and demonstrated the value of a further consideration of predictive power for HRM research. We illustrated the predictive power assessments by using our HPWS example, which demonstrated predictive power in addition to explanatory power. Likewise, we demonstrated model comparisons between the two modelling options of our introduced

HPWSs. In our analysis, the HCM did not significantly outperform the model without a HOC. In our case, this may be a result of the rather simple model that we chose for illustrative purposes, and should not be over-interpreted for the field of HRM systems research. We encourage researchers to use the HCM, and to use the predictive power assessments that PLS-SEM offers to enrich their model evaluations. This also applies to predictive model comparisons when there are alternative theoretically established models. Overall, this contributes to the selection of models that not only demonstrate robustness and generalizability across diverse HRM scenarios, but also predict outcomes accurately.

Researchers who are interested in developing formative measurements and HCM in their empirical study designs can consult further guidelines to develop scales (MacKenzie, Posdakoff & Podsakoff, 2011), in addition to core readings in PLS-SEM (Hair et al., 2022; Hair, Sarstedt, Ringle, et al., 2024). Also, while we illustrated our ideas using a HPWS example, further illustrations and guidance in HRM are available (e.g. Legate, Hair, et al., 2023; Legate, Ringle, et al., 2023; Gubernator et al., 2024).

In summary, we have highlighted the importance of latent constructs with formative measurement models, HCMs, and predictive model assessments in HRM systems research. By illustrating modeling options in PLS-SEM, providing current guidelines, and their application, we hope to encourage SHRM researchers to make greater use of them in order to improve the conceptual accuracy, methodological rigor, and practical relevance of HRM systems research.

Notes

1. Boon et al. (2019) summarized all these approaches under additive approaches.
2. Jiang et al. (2012) distinguish between HRM policies and HRM practices. According to the authors, HRM policies represent "choices" and "guidelines" (e.g. pay for performance); HRM practices are different ways of implementing the HRM policies (e.g. pay for performance is achieved via a piece-rate system or stock option plans). SHRM researchers mostly focus on HRM policies, but use the term HRM practices. We do the same here, not distinguishing further between HRM policies and HRM practices.
3. Convergent validity is another criterion with which to assess formative measurement models. Redundancy analyses are recommended for statistically evaluating convergent validity. In these analyses, researchers assess whether the correlation of the construct with an alternative measure of the same concept is 0.70 or higher. This alternative measure needs to be integrated into the research design. It could be an alternative reflectively measured construct or a (global) single-item covering the construct's essence (Hair et al., 2022). The HRM systems literature has not yet developed such an alternative measure for an HRM system, which may be a challenging endeavor given most HRM systems' comprehensive nature. Therefore, although we do not consider this criterion here, we call for further research on this.

Disclosure statement

This research uses the statistical software SmartPLS (https://www.smartpls.com). Christian M. Ringle acknowledges a financial interest in SmartPLS.

Data availability statement

The data supporting this study's findings are available from the corresponding author on request.

References

Aguirre-Urreta, M. I., & Rönkkö, M, Florida International University. (2018). Statistical inference with PLSc using bootstrap confidence intervals. *MIS Quarterly*, *42*(3), 1001–1020. https://doi.org/10.25300/MISQ/2018/13587

Alothmany, R., Jiang, Z., & Manoharan, A. (2023). Linking high-performance work systems to affective commitment, job satisfaction, and career satisfaction: Thriving as a mediator and wasta as a moderator. *The International Journal of Human Resource Management*, *34*(19), 3787–3824. https://doi.org/10.1080/09585192.2022.2157681

Appelbaum, E., Bailey, T., Berg, P., & Kalleberg, A. L. (2000). *Manufacturing advantage: Why high performance work systems pay off*. Cornell Univ. Press.

Becker, J.-M., Cheah, J.-H., Gholamzade, R., Ringle, C. M., & Sarstedt, M. (2023). PLS-SEM's most wanted guidance. *International Journal of Contemporary Hospitality Management*, *35*(1), 321–346. https://doi.org/10.1108/IJCHM-04-2022-0474

Beijer, S., Peccei, R., Van Veldhoven, M., & Paauwe, J. (2021). The turn to employees in the measurement of human resource practices: A critical review and proposed way forward. *Human Resource Management Journal*, *31*(1), 1–17. https://doi.org/10.1111/1748-8583.12229

Bollen, K. A., & Diamantopoulos, A. (2017). In defense of causal-formative indicators: A minority report. *Psychological Methods*, *22*(3), 581–596. https://doi.org/10.1037/met0000056

Boon, C., Den Hartog, D. N., & Lepak, D. P. (2019). A systematic review of human resource management systems and their measurement. *Journal of Management*, *45*(6), 2498–2537. https://doi.org/10.1177/0149206318818718

Bos-Nehles, A., Townsend, K., Cafferkey, K., & Trullen, J. (2023). Examining the ability, motivation and opportunity (AMO) framework in HRM research: Conceptualization, measurement and interactions. *International Journal of Management Reviews*, *25*(4), 725–739. https://doi.org/10.1111/ijmr.12332

Cenfetelli, R.T. and Bassellier, G. (2009), Interpretation of formative measurement in information systems research, *MIS Quarterly*, 4, *33*, 689–708. https://doi.org/10.2307/20650323

Chin, W., Cheah, J.-H., Liu, Y., Ting, H., Lim, X.-J., & Cham, T. H. (2020). Demystifying the role of causal-predictive modeling using partial least squares structural equation modeling in information systems research. *Industrial Management & Data Systems*, *120*(12), 2161–2209. https://doi.org/10.1108/IMDS-10-2019-0529

Cooke, F. L., Xiao, M., & Chen, Y. (2021). Still in search of strategic human resource management? A review and suggestions for future research with China as an example. *Human Resource Management*, *60*(1), 89–118. https://doi.org/10.1002/hrm.22029

Danks, N. P., Sharma, P. N., & Sarstedt, M. (2020). Model selection uncertainty and multimodel inference in partial least squares structural equation modeling (PLS-SEM). *Journal of Business Research*, *113*, 13–24. https://doi.org/10.1016/j.jbusres.2020.03.019

Diamantopoulos, A., & Riefler, P. (2011). Using formative measures in international marketing models: A cautionary tale using consumer animosity as an example. *Advances in International Marketing*, *22*, 11–30.

Diamantopoulos, A., Riefler, P., & Roth, K. P. (2008). Advancing formative measurement models. *Journal of Business Research*, *61*(12), 1203–1218. https://doi.org/10.1016/j.jbusres.2008.01.009

Diamantopoulos, A., Sarstedt, M., Fuchs, C., Wilczynski, P., & Kaiser, S. (2012). Guidelines for choosing between multi-item and single-item scales for construct measurement: A predictive validity perspective. *Journal of the Academy of Marketing Science*, *40*(3), 434–449. https://doi.org/10.1007/s11747-011-0300-3

Fornell, C., & Bookstein, F. L. (1982). Two structural equation models: LISREL and PLS applied to consumer exit-voice theory. *Journal of Marketing Research*, *19*(4), 440–452. https://doi.org/10.1177/002224378201900406

Garmendia, A., Elorza, U., Aritzeta, A., & Madinabeitia-Olabarria, D. (2021). High-involvement, job satisfaction and productivity: A two wave longitudinal study of a Spanish retail company. *Human Resource Management Journal*, *31*(1), 341–357. https://doi.org/10.1111/1748-8583.12307

Gubernator, P., Hauff, S., & Günther, N. (2024). The effectiveness of well-being-oriented human resource management in the context of telework. *The International Journal of Human Resource Management*, 1–29. https://doi.org/10.1080/09585192.2024.2354829

Guest, D. E. (2017). Human resource management and employee well-being: Towards a new analytic framework. *Human Resource Management Journal*, *27*(1), 22–38. https://doi.org/10.1111/1748-8583.12139

Guest, D. E. (2025). Strengthening links between HRM theories, HR practices and outcomes: A proposal to advance research on HRM and outcomes. *Human Resource Management Journal*, *35*(1), 319–335. https://doi.org/10.1111/1748-8583.12569

Hair, J. F., Hult, T. M., Ringle, C. M., & Sarstedt, M. (2022). *A primer on partial least squares structural equation modeling (PLS-SEM) (3rd ed.)*. Sage.

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, *31*(1), 2–24. https://doi.org/10.1108/EBR-11-2018-0203

Hair, J. F., Sarstedt, M., Ringle, C. M., & Gudergan, S. P. (2024). *Advanced issues in partial least squares structural equation modeling (PLS-SEM) (2nd ed.)*. Sage.

Hair, J. F., Sharma, P. N., Sarstedt, M., Ringle, C. M., & Liengaard, B. D. (2024). The shortcomings of equal weights estimation and the composite equivalence index in PLS-SEM. *European Journal of Marketing*, *58*(13), 30–55. https://doi.org/10.1108/EJM-04-2023-0307

Hauff, S. (2021). Analytical strategies in HRM systems research: A comparative analysis and some recommendations. *The International Journal of Human Resource Management*, *32*(9), 1923–1952. https://doi.org/10.1080/09585192.2018.1547779

Jackson, S. E., Schuler, R. S., & Jiang, K. (2014). An aspirational framework for strategic human resource management. *Academy of Management Annals*, *8*(1), 1–56. https://doi.org/10.5465/19416520.2014.872335

Jiang, K., & Messersmith, J. (2018). On the shoulders of giants: A meta-review of strategic human resource management. *The International Journal of Human Resource Management*, *29*(1), 6–33. https://doi.org/10.1080/09585192.2017.1384930

Jiang, K., Lepak, D. P., Han, K., Hong, Y., Kim, A., & Winkler, A.-L. (2012a). Clarifying the construct of human resource systems: Relating human resource management to employee performance. *Human Resource Management Review*, *22*(2), 73–85. https://doi.org/10.1016/j.hrmr.2011.11.005

Jiang, K., Lepak, D. P., Hu, J., & Baer, J. C. (2012b). How does human resource management influence organizational outcomes? A meta-analytic investigation of mediating mechanisms. *Academy of Management Journal*, *55*(6), 1264–1294. https://doi.org/10.5465/amj.2011.0088

Jöreskog, K. G. (1982). The LISREL approach to causal model-building in the social sciences. In K.G. Jöreskog and H. Wold (Eds.), *Systems under indirect observations: Part II*. (pp. 81–100). North-Holland.

Jöreskog, K. G., & Wold, H. (1982). The ML and PLS techniques for modeling with latent variables: Historical and comparative aspects. In H. Wold and K.G. (Eds.), *Systems under indirect observation, Part I*. North-Holland. pp. 263–270.

Kaufman, B. E. (2015). Evolution of strategic HRM as seen through two founding books: A 30th anniversary perspective on development of the field. *Human Resource Management*, *54*(3), 389–407. https://doi.org/10.1002/hrm.21720

Kaufman, B. E. (2020). The real problem: The deadly combination of psychologisation, scientism, and normative promotionalism takes strategic human resource management down a 30-year dead end. *Human Resource Management Journal*, *30*(1), 49–72. https://doi.org/10.1111/1748-8583.12278

Kim, J. (2024). Building cooperative and high-performance organizations with high-performance work systems: The role of firm age and industry dynamism. *The International Journal of Human Resource Management*, *35*(13), 2222–2258. https://doi.org/10.1080/09585192.2024.2330983

Legate, A. E., Hair, J. F., Chretien, J. L., & Risher, J. J. (2023). PLS-SEM: Prediction-oriented solutions for HRD researchers. *Human Resource Development Quarterly*, *34*(1), 91–109. https://doi.org/10.1002/hrdq.21466

Legate, A. E., Ringle, C. M., & Hair, J. F. (2023). PLS-SEM: A method demonstration in the R statistical environment. *Human Resource Development Quarterly*, *35*(4), 501–529. https://doi.org/10.1002/hrdq.21517

Lepak, D. P., Liao, H., Chung, Y., & Harden, E. E. (2006). A conceptual review of human resource management systems in strategic human resource management research. *Research in Personnel and Human Resource Management*, *25*, 217–271.

Liengaard, B. D., Sharma, P. N., Hult, G. T. M., Jensen, M. B., Sarstedt, M., Hair, J. F., & Ringle, C. M. (2021). Prediction: coveted, yet Forsaken? Introducing a cross-validated predictive ability test in partial least squares path modeling. *Decision Sciences*, *52*(2), 362–392. https://doi.org/10.1111/deci.12445

Lohmöller, J.-B. (1989). *Latent variable path modeling with partial least squares*., Physica-Verl.

MacKenzie, S.B., Podsakoff, P.M., Podsakoff, N.P. (2011), Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques, *MIS Quarterly*, 2, *35*, 293–334. https://doi.org/10.2307/23044045

MacKenzie, S. B., Podsakoff, P. M., & Jarvis, C. B. (2005). The problem of measurement model misspecification in behavioral and organizational research and some recommended solutions. *The Journal of Applied Psychology*, *90*(4), 710–730. https://doi.org/10.1037/0021-9010.90.4.710

McNeish, D. (2023). Psychometric properties of sum scores and factor scores differ even when their correlation is 0.98: A response to Widaman and Revelle. *Behavior Research Methods*, *55*(8), 4269–4290. https://doi.org/10.3758/s13428-022-02016-x

McNeish, D., & Wolf, M. G. (2020). Thinking twice about sum scores. *Behavior Research Methods*, *52*(6), 2287–2305. https://doi.org/10.3758/s13428-020-01398-0

Ogbonnaya, C., & Messersmith, J. (2019). Employee performance, well-being, and differential effects of human resource management subdimensions: Mutual gains or conflicting outcomes? *Human Resource Management Journal*, *29*(3), 509–526. https://doi.org/10.1111/1748-8583.12203

Petter, S., & Hadavi, Y. (2023). Use of partial least squares path modeling within and across business disciplines. In H. Latan, J.F. Hair, and R. Noonan (Eds.), *Partial least squares path modeling: Basic concepts, methodological issues and applications* (pp. 55–79) Springer International Publishing.

Posthuma, R. A., Campion, M. C., Masimova, M., & Campion, M. A. (2013). A high performance work practices taxonomy: Integrating the literature and directing future research. *Journal of Management*, *39*(5), 1184–1220. https://doi.org/10.1177/0149206313478184

Richter, N. F., & Tudoran, A. A. (2024). Elevating theoretical insight and predictive accuracy in business research: Combining PLS-SEM and selected machine learning algorithms. *Journal of Business Research*, *173*, 114453. https://doi.org/10.1016/j.jbusres.2023.114453

Rigdon, E. E., Sarstedt, M., & Ringle, C. M. (2017). On comparing results from CB-SEM and PLS-SEM. Five perspectives and five recommendations. *Marketing ZFP*, *39*(3), 4–16. https://doi.org/10.15358/0344-1369-2017-3-4

Rigdon, E., Sarstedt, M., & Moisescu, O. I. (2023). Quantifying model selection uncertainty via bootstrapping and Akaike weights. *International Journal of Consumer Studies*, *47*(4), 1596–1608. https://doi.org/10.1111/ijcs.12906

Ringle, C. M., Sarstedt, M., Mitchell, R., & Gudergan, S. P. (2020). Partial least squares structural equation modeling in HRM research. *The International Journal of Human Resource Management*, *31*(12), 1617–1643. https://doi.org/10.1080/09585192.2017.1416655

Ringle, C. M., Wende, S., & Becker, J.-M. (2024). *SmartPLS 4*. SmartPLS.

Sarstedt, M., & Danks, N. P. (2022). Prediction in HRM research – A gap between rhetoric and reality. *Human Resource Management Journal*, *32*(2), 485–513. https://doi.org/10.1111/1748-8583.12400

Sarstedt, M., Adler, S. J., Ringle, C. M., Cho, G., Diamantopoulos, A., Hwang, H., & Liengaard, B. D. (2024). Same model, same data, but different outcomes: Evaluating the impact of method choices in structural equation modeling. *Journal of Product Innovation Management*, *41*(6), 1100–1117. https://doi.org/10.1111/jpim.12738

Sarstedt, M., Hair, J. F., Cheah, J.-H., Becker, J.-M., & Ringle, C. M. (2019). How to specify, estimate, and validate higher-order constructs in PLS-SEM. *Australasian Marketing Journal*, *27*(3), 197–211. https://doi.org/10.1016/j.ausmj.2019.05.003

Sarstedt, M., Hair, J. F., Pick, M., Liengaard, B. D., Radomir, L., & Ringle, C. M. (2022). Progress in partial least squares structural equation modeling use in marketing research in the last decade. *Psychology & Marketing*, *39*(5), 1035–1064. https://doi.org/10.1002/mar.21640

Sarstedt, M., Hair, J. F., Ringle, C. M., Thiele, K. O., & Gudergan, S. P. (2016). Estimation issues with PLS and CBSEM: Where the bias lies!. *Journal of Business Research*, *69*(10), 3998–4010. https://doi.org/10.1016/j.jbusres.2016.06.007

Sarstedt, M., Ringle, C. M., & Hair, J. F. (2021). Partial least squares structural equation modeling. In C. Homburg, M. Klarmann, and A.E. Vomberg (Eds.), *Handbook of market research* (pp. 1–47). Springer.

Sharma, P. N., Liengaard, B. D., Hair, J. F., Sarstedt, M., & Ringle, C. M. (2023). Predictive model assessment and selection in composite-based modeling using PLS-SEM: Extensions and guidelines for using CVPAT. *European Journal of Marketing*, *57*(6), 1662–1677. https://doi.org/10.1108/EJM-08-2020-0636

Sharma, P. N., Sarstedt, M., Ringle, C. M., Cheah, J.-H., Herfurth, A., & Hair, J. F. (2024). A framework for enhancing the replicability of behavioral MIS research using prediction oriented techniques. *International Journal of Information Management*, *78*, 102805. https://doi.org/10.1016/j.ijinfomgt.2024.102805

Sharma, P. N., Sarstedt, M., Shmueli, G., Kim, K. H., & Thiele, K. O. (2019). PLS-based model selection: The role of alternative explanations in information systems research. *Journal of the Association for Information Systems*, *40*, 346–397. https://doi.org/10.17705/1jais.00538

Shmueli, G. (2010). To explain or to predict? *Statistical Science*, *25*(3), 289–310. https://doi.org/10.1214/10-STS330

Shmueli, G., Ray, S., Velasquez Estrada, J. M., & Chatla, S. B. (2016). The elephant in the room: Evaluating the predictive performance of PLS models. *Journal of Business Research*, *69*(10), 4552–4564. https://doi.org/10.1016/j.jbusres.2016.03.049

Shmueli, G., Sarstedt, M., Hair, J. F., Cheah, J., Ting, H., Vaithilingam, S., & Ringle, C. M. (2019). Predictive model assessment in PLS-SEM: Guidelines for using PLSpredict. *European Journal of Marketing*, *53*(11), 2322–2347. https://doi.org/10.1108/EJM-02-2019-0189

Wold, H. (1982). Soft modeling: The basic design and some extensions. In K.G. Jöreskog and H. Wold (Eds.), *Systems under indirect observations: Part II* (pp. 1–54). North-Holland.

Zhou, Q., Edafioghor, T. E., Wu, C. H., & Doherty, B. (2023). Building organisational resilience capability in small and medium-sized enterprises: The role of high-performance work systems. *Human Resource Management Journal*, *33*(4), 806–827. https://doi.org/10.1111/1748-8583.12479