

# Cross-Country Analysis of HRM Parameters in Emerging Markets: An Assessment of Measurement Invariance

Tamer K. Darwish <sup>1,2</sup>, Satwinder Singh,<sup>3</sup> Georgios Batsakis <sup>4,5</sup>  
and Kristina Potočnik<sup>6</sup>

<sup>1</sup>Business School, University of Gloucestershire, GL2 9HW, Gloucester, UK, <sup>2</sup>Gulf Financial Center, Gulf University for Science and Technology, 32093, Hawally, Kuwait, <sup>3</sup>Business School, University of Dubai, Dubai, United Arab Emirates, <sup>4</sup>Alba Graduate Business School, The American College of Greece, 115 28, Athens, Greece, <sup>5</sup>Brunel Business School, Brunel University London, UB8 3PH, Uxbridge, UK, and <sup>6</sup>Business School, University of Edinburgh, EH8 9JS, Edinburgh, UK  
Corresponding author email: [tdarwish@glos.ac.uk](mailto:tdarwish@glos.ac.uk)

**In this paper we aim to critically discuss the challenges and benefits of using survey instruments (SIs) by measuring human resource management (HRM) parameters across five emerging markets – Brunei Darussalam, India, Jordan, Kingdom of Saudi Arabia and United Arab Emirates. In so doing, we proceed to an assessment of the measurement invariance of HRM parameters measured by SIs across these five countries. Based on our field experience and empirical results showing lack of measurement invariance across these five national contexts, we provide recommendations for conducting comparative analysis of HRM parameters in emerging markets, ultimately contributing to the HRM convergence–divergence debate in this type of national context.**

## Introduction

Survey instruments (SIs) – or data collection tools whereby each respondent answers the same set of questions or items in a pre-specified order – are a frequently used tool in management research in general, and in comparative management and human resource management (HRM) research in particular (Akinci and Saunders, 2015; Collinson and Pettigrew, 2009; Robinson, 2018). The widespread use of SIs in comparative management research, which focuses on analysing differences and similarities between employees, managers, workgroups, firms, subsidiaries and their headquarters, amongst others, is not surprising (Collinson and Pettigrew, 2009). Namely, SIs enable data collection on larger samples in multiple locations or countries, and generate data for statistical analyses to establish the magnitude of any

differences or similarities between the units under investigation. In case of comparative HRM research, which often explores cross-country differences, recent systematic reviews point to a steady increase in quantitative research with surveys being the most frequently used method of data collection (Cooke, Veen and Wood, 2017, 2019). For instance, some of the most known large-scale and ongoing international surveys in the comparative HRM area are the Cranfield project on HRM practices in Europe and beyond – that is, the CRANET project (Brewster, Mayrhofer and Reichel, 2011) and the global leadership and organizational effectiveness project – that is, the GLOBE project (House *et al.*, 2013). While these projects started off focusing on developed countries, most of them covered data collection in developing countries, often called emerging markets, as well.

Given this increasingly widespread use of SIs in comparative HRM literature, our aim in this paper is to critically discuss, analyse and review the benefits and challenges of using surveys in studying HRM parameters across different emerging markets and provide recommendations for survey design based on our field experience and unique findings on cross-country comparability of HRM parameters in this type of national context. Our study's empirical setting focuses on emerging markets in general and five specific emerging markets in particular (i.e. Brunei Darussalam, India, Jordan, Kingdom of Saudi Arabia – KSA and United Arab Emirates – UAE). We do so for the following reasons. First, we lack knowledge on how comparable HRM surveys are in the context of emerging markets, which is an important omission given the increasing interest in studying HRM in these national contexts (Brewster, Mayrhofer and Cooke, 2015; Darwish *et al.*, 2020). Second, by focusing only on countries that primarily share cultural and other socio-political similarities, and secondarily a similar economic development and outlook, we address the call for more research on the HRM divergence–convergence debate using a relatively homogenous sample (Brewster, Mayrhofer and Cooke, 2015; Malik, Pereira and Budhwar, 2022). Third, our decision to focus on the above five emerging markets is also based on the relatively moderate to strong institutional setting of these countries compared to other emerging markets. Among others, a crucial factor that shapes the extent of adoption of HRM practices in firms is the relative strength of institutions in a given country (Dibben *et al.*, 2017). Our sample's firms are sufficiently homogenized, and relatively Westernized in terms of the extent of institutional quality, as this is captured by six indicators, namely voice and accountability, political stability and absence of violence, government effectiveness, regulatory quality, rule of law and control of corruption (Kaufmann, Kraay and Mastruzzi, 2009).<sup>1</sup>

In sum, this paper addresses the call for enhanced rigour in HRM measurement, which is necessary for robust and meaningful theory development in comparative HRM (Boon, Den Hartog and Lepak, 2019). Our unique findings around the cross-country comparability of SIs to measure di-

verse HRM parameters in five emerging markets advance the literature on measurement methods in comparative HRM by showing how appropriate it might be to measure HRM parameters using the same SI in countries that may appear similar at first sight – that is, Asian emerging markets. While there is extant research in the field of comparative HRM showing that HR practices applied in one country are not readily transferable to another country or cultural context (e.g. Haak-Saheem and Darwish, 2021; Rowley, Quang and Warner, 2007), we focus on emerging markets firms only, thus providing answers on the potential divergence versus convergence of HRM practices within a seemingly homogenous sample.

The focus on emerging markets reveals that this group is not as homogenous as one would expect. While there is a consensus among HRM scholars that emerging markets are highly similar in their context and directional convergence (i.e. the same factors affect both the nature of HRM and the challenges facing the practice of HRM) (Budhwar, Varma and Patel, 2016), our findings do not confirm this view. Thus, our recommendations for HRM survey design in emerging markets can further advance the debate on HRM convergence–divergence in emerging markets (Brewster, Mayrhofer and Cooke, 2015). Also, compared to the vast majority of studies in the field of comparative HRM, we tested the measurement invariance of key HRM parameters using data collected over an extensive time period (2010–2018), adding a longitudinal dimension to the cross-country comparison. This methodological idiosyncrasy – that is, the fact that data were collected over an extensive time period – contributes towards realizing the need to account not only for the contextual (dis)similarities, but also for the temporal dimension affecting the adoption of HR practices across explored emerging markets (Hoorani, Plakoyiannaki and Gibbert, 2023). As such, we recommend that more emphasis should be put on better utilizing temporal effects in investigating various HRM practices across different markets. Finally, our recommendations are grounded in the assessment of measurement invariance on a complete set of HRM-related constructs linked to HR director role, recruitment, training and retention, appraisals, incentives and rewards. The focus on a more inclusive use and assessment of HRM parameters makes our recommendations more comprehensive.

<sup>1</sup>A detailed analysis of the institutional background and overall evaluation of these five countries is reported in the online Supporting Information.

Our paper starts by providing an overview of SIs as a method of field research in the domestic and international context of HRM, followed by a critical discussion of the benefits and challenges of using SIs in each of the five emerging markets of data collection. We then assess the measurement invariance of our HRM survey data from these five countries to provide unique empirical findings of comparability of SIs assessing different HRM parameters across these national contexts. Grounded in these findings, we conclude with a set of recommendations for survey design in emerging markets.

### HRM surveys in emerging markets

Emerging markets have most frequently been defined to represent those countries 'whose national economies have grown rapidly, where industries have undergone and are continuing to undergo dramatic structural changes, and whose markets hold promise despite volatile and weak legal systems' (Luo and Tung, 2007, p. 482). Most of the academic research on HRM in emerging markets has focused on studying multinational enterprises (MNEs) with subsidiaries in emerging markets and HRM issues, in mergers and acquisitions. Another angle to this research comes from studies that focus on MNEs that are headquartered in emerging markets, such as China and India, with subsidiaries in the West (Skuzza, McDonnell and Scullion, 2015). The use of SIs in such contexts might be the only method of data collection that allows researchers to collect adequate data in a reasonable period to learn about, and compare, HRM issues between the headquarters and subsidiaries, or to understand how such issues in organizations from emerging markets compare with organizations in developed countries. Some studies have explored HRM issues in local organizations located in emerging markets that have no ties to the international corporate world. All of these studies have the potential to make a significant contribution to the literature and practice, as one of the major challenges of MNEs is to understand what HRM practices would work effectively in different cultural contexts, which would ultimately help them with their competitive advantage in global markets (Cooke, Veen and Wood, 2017). Similarly, domestic organizations from emerging markets are keen to understand what HRM

practices that were proved to work in developed countries could also work in their context.

Most of the existing research that has addressed these questions in emerging markets, by means of SIs, has been designed for the purpose of a particular study or context. It is beyond the scope of this paper to provide a systematic literature review of these studies as they have recently been systematically reviewed elsewhere (see Cooke, Veen and Wood, 2017, 2019, 2020). However, we briefly review illustrative papers from different emerging markets that were published in the last 5 years and highlight key large international surveys that provide comprehensive HRM data from different emerging and developed markets using identical or near identical surveys.

In the Chinese context, Yang *et al.* (2021) recently explored the joint role of more traditional Chinese *guanxi* HRM practices and Western individual pay-for-performance practices on emotional exhaustion and job performance. They collected their data in multiple groups from one privately owned company, using designated contacts in each group that distributed the questionnaires in person. They used previously designed scales to measure their core concepts. Ge and Zhao (2020) explored the effectiveness of a hybrid HRM system of foreign-invested enterprises (FIEs) in China. They collected the data by means of SIs which they designed for the purposes of the study, based on their theoretical arguments of hybrid systems, as well as on some established scales and indicators (e.g. cultural differences were operationalized by means of Hofstede's, 1980 index). Gao *et al.* (2020) explored how the improvement in career development opportunities may worsen the experience of emotional exhaustion, particularly for those knowledge workers who experienced the best improvement in their career opportunities. Similar to Yang *et al.* (2021), they constructed their questionnaire by using items from previously validated and established scales.

In the Indian context, Datta *et al.* (2023) explored different facets of talent development climate and how these facets may mediate the relationships between employee perceived HRM practices and innovative work behaviour, respectively. Their survey was composed of their own talent development climate questionnaire based on their findings from a qualitative study, as well as previously validated scales. Compared to this study, Mariappanadar (2020) focused on Indian

domestic organizations to explore the bundles of motivation-enhancing HRM practices. The data for this study were collected by means of online surveys using previously validated scales. More recently, Lakshman *et al.* (2022) explored the relationships between flexibility-oriented HRM practices and innovation, as mediated by intellectual capital. They used a sample of HR and other senior managers from 254 Chinese organizations and their 143 counterparts from India. They used identical (and previously established) scales to measure their core variables in both samples. This is one of the few studies that used primary data collection in more than one emerging market to explore their hypotheses. In this case, the authors found a large degree of convergence in their results between China and India.

In the context of Russia, Soklov and Zavyalova (2021) recently discerned the ability, motivation and opportunity-enhancing dimensions of HRM systems to explore their relationships with human, social and structural capital, respectively. Rather than using an online or paper-and-pencil survey, they collected their data by means of a telephone survey, composed of established and previously validated scales, on a sample of 184 HR and other senior managers working in Russian knowledge-intensive industries. Another insightful study from Russia explored the relationships between talent management practices and competitive advantage in companies that are headquartered in Russia (Latukha, 2018). In this case, the HR managers were asked to complete the survey either in online or paper-and-pencil format and their findings showed that succession planning and career development were particularly important to establishing competitive advantage in Russia.

In addition to these studies, there are various large international surveys that provide comprehensive data on HRM practices from different emerging and developed markets using identical or near identical surveys. For instance, the Global Leadership and Organizational Effectiveness (GLOBE) project is also an example of a large multi-country research endeavour that focused on data collection by means of SIs (House *et al.*, 2013). The main aim of this project is to explore the cross-cultural differences in organizational practices and conceptions of leadership. A total of 62 developed and emerging countries have been taking part in the project and, in each country, the data collection has been done in three local

organizations representing the financial services, food processing and telecommunications sectors. The CRANET project (Brewster, Mayrhofer and Reichel, 2011) is another example of a long-standing comprehensive international project that investigates organizational policies and practices in people management by means of SIs. Coordinated by Cranfield School of Management, the project was initially launched in five European countries, but the network has grown enormously in recent years and now covers more than 35 countries, including emerging markets such as Brazil, Russia, India, China and South Africa. These large international surveys have led to a number of publications across the last three decades.

### **Key challenges in replicating HRM SIs across countries: What do we know so far?**

Business schools' doctoral training programmes cover a range of best practices about how to construct SIs and how to execute them. These include a wide array of aspects, such as principles of developing SIs to be executed in single and multi-country contexts, how to ensure uniform sampling techniques across countries, how to achieve acceptable survey response rates, whether to use partner organizations for data collection, how to manage multi-team international research projects (e.g. GLOBE involved 170 researchers at some point), how to improve the level of consistency in relation to the reporting of the sample frame, what is considered to be an acceptable sample size and useful responses (Cooke *et al.*, 2019), amongst others. This is what we can term an 'idealized state' or 'starting point' in the learning process, as opposed to what researchers face in the field, and what we can term 'realities of applied survey research on the ground' (see e.g. Zahra, 2011). The 'idealized state' of affairs that we academics teach in the classroom largely refers to conducting surveys in the developed world, where the socio-economic system is much more organized, as opposed to emerging markets where institutions are still evolving and where a mixture of political, economic, legal and social factors can play a significant role when scholars seek to conduct their research (Haak-Saheem, Festing and Darwish, 2017; Schotter, Meyer and Wood, 2021). In some emerging markets, access to participants might be

difficult online and data collection is therefore still done by means of paper-and-pencil questionnaires (e.g. Yang *et al.*, 2021) or telephone interviews (e.g. Soklov and Zavyalova, 2021). Inter-country and intra-country cross-cultural issues may reign supreme in these countries, causing considerable challenges to researcher(s). We next discuss our experience in conducting SIs across Brunei Darussalam, India, Jordan, KSA and UAE.

#### *Five emerging markets under the microscope*

*Jordan.* At the time of survey administration, Jordan was grappling with uncertainty in the region, high unemployment, slow economic growth, low job security, migration and brain drain. Clearly, conducting research and collecting data from a region such as this proved very challenging. Managers were not interested in participating in our research, given that this was a low priority for them. Still, we tried to follow the normal ‘idealized route’ by drafting a comprehensive instrument on HRM and sending it out by post, following up with reminders and so on. We then changed technique and tried to follow the online route – but again with no success. For cultural reasons, it seemed that following a face-to-face approach was rather more effective than other methods (e.g. online) to collect the data, and also ensured a higher response rate. So it transpired that our approach must be altered. The decision-making process was twofold. First, we cut down the target population from multi-sectors to one (financial sector), where the population size was small. A sample of just below 100 companies operating in this sector was reached. Second, we decided to approach all HR managers in the population in person. This approach proved successful, though it took about 6 months from start (pilot trial) to finish. A positive of this approach turned out to be that we got to learn a lot more about the sector and the economy in general, in addition to gaining first-hand insights into the HR-related issues facing the economy and the concerned firm where the data were being collected.

*KSA and UAE.* A relatively reliable response rate was achieved in these two countries (a sample of 255 companies for KSA and 127 companies for UAE across all sectors). Although subtle differences remain among Arab nations, the KSA and UAE are somewhat similar in some of their cul-

tural and institutional aspects. However, within these contexts, getting access to collect data from HR directors posed a separate set of problems. For instance, HR directors were not only Emiratis or Saudi nationals, but many were expats who had their own issues (sometimes insecurities) with the release of information. Also, unlike Jordan, target firms in the KSA and UAE were scattered across the country, which meant that it would be an uphill task to conduct face-to-face interviews with a view to complete the SIs. To overcome these difficulties, it was decided to invoke personal connections in ministries so that we could obtain a good covering letter to go with the SIs. It paid off to invest in personal connections and contacts (i.e. *wasta*),<sup>2</sup> as such letters could be one of the most important conduits to gain access to data in the context of the Middle East. *Wasta* refers to the powerful informal social institutions in the day-to-day activities and operations of businesses conducted in the region (e.g. Haak-Saheem and Darwish, 2021).

In the Middle East in general, and in the Gulf region in particular, individuals and organizations have their own unique organizational cultures, governance and structures, which are different from the rest of the world (e.g. Haak-Saheem and Darwish, 2021; Zahra, 2011). It is the understanding of these organizational, cultural and governance issues that could be crucial in helping researchers to collect relevant data. One issue which was confirmed by the research team was that trust and long-term relationship building was a block in the national and organizational cultures of the Middle East (Haak-Saheem, Festing and Darwish, 2017; Zahra, 2011). Understanding and voicing the perspectives and values of those managers and professionals who work in the targeted companies is indeed essential within such cultures. The research team also confirmed that it is essential for targeted managers to fully understand the purpose and use of research, and even engage with all its stages to increase the chances of participation and ensure more reliable findings. This was very much our case in the KSA and UAE, where many managers were pleased to participate in our research activities in a reciprocal way where we also shared our findings with them. Our experience also tells us that establishing research teams which have collaborations with local scholars could be helpful in

<sup>2</sup> *Guanxi* in China, *jeitinho* in Brazil, *svyazi* and *blat* in Russia are very similar concepts to *wasta*.

getting access to data and stimulating interest among local managers and companies.

*Brunei Darussalam.* Executing SIs in Brunei posed yet another set of issues. We first tried to locate a list of companies, but to our surprise we could not find a reliable list. After some deliberation, it was decided that for the purpose of proper sampling we needed to have knowledge of the population. So, as a start, this list had to be laboriously compiled with help from the Ministry of Finance, the Ministry of Industry and Primary Resources and the Brunei Economic Development Board. We identified a population of 465 firms, which was the starting point of the project. Based on this list, a sample size of 214 was then drawn randomly for the purpose of the study. The survey was to be addressed to HR directors, but we found that some small organizations did not have a HR director. However, all companies contacted had someone in charge of personnel. In larger companies, they were called the HR director or HR manager. In smaller companies, those in charge of personnel issues were knowledgeable about the affairs of the company, although they did not hold the title of HR director or HR manager. Often, we found that in smaller companies the interface with managers was candid and fruitful.

Owing to the environment of the country being explored, it was decided that the questionnaire would be distributed in English as well as Malay, unlike for other countries where the survey remained in English. In international research, translation is important, especially if the questions are to have identical meaning to all participants. Back translation is the most commonly used method in multi-country research (e.g. House *et al.*, 2013); however, the limitations associated with employing the back-translation approach have led to this avenue being discarded. These limitations include possible differences in language use between bi- and mono-lingual speakers (House *et al.*, 2013), as well as back translation only providing a literal translation from one language to another, with the translation not capturing the intended sense of statement (Douglas and Craig, 2007). The parallel translation approach – which has been advocated as the preferred method for achieving equivalence in meaning (Hambleton, 1993) – was selected to translate the questionnaire from English to Malay.

Considering the four sources of error, which can lead to unacceptable survey errors, namely sampling error, non-coverage error, non-response error and measurement error (Groves, 1990), we implemented the total design method as suggested by Dillman (1978). Non-response error occurs when some members do not respond to the questionnaire. This was potentially alleviated by ensuring a respondent-friendly questionnaire, which progressed through several drafts, as well as undergoing a pilot study. To make it easier for organizations, the cover notes showed that personalization of correspondence and return envelopes were provided. Also, there were multiple contacts with the desired organizations, following the direction specified by the total design method. Measurement error tends to arise from the characteristics of the respondents or the characteristics of the questions. This error was minimized through directing the questionnaire to the HR directors of the organizations, who would have knowledge of the HRM activities in their organizations, and by evaluating the questionnaire design, which had already been used in previous research several times to ensure a fit with the focus of this research before undergoing the pilot study process.

*India.* Executing our SI in India began with a search of the database of companies from which to choose a sample. This posed issues, as the list we were looking for was not available from one source, and compiling it from various sources<sup>3</sup> soon proved to be a major challenge. Given the size of the country, firms were scattered all over and it would be a near impossible task to visit them in person. Obtaining adequate and reliable data was the main concern, so we decided that a marketing organization should be entrusted with this task. We wanted the data to be collected from a cross-section of firms located across six centres (Delhi, Calcutta, Bombay, Chennai, Bangalore, Indore/Ahmedabad). The contracted company<sup>4</sup> succeeded in compiling a dataset of 300 such enterprises, which became our starting point. We approached all 300 companies in writing, following up with telephone calls and in some instances seeing them in person. A pilot study was first

<sup>3</sup>Department of Company Affairs, Ministry of Finance, trade and other sources.

<sup>4</sup>The contracted company, Synovate Comcon, is part of the international research network Ipsos, which is in the top three leading market research companies worldwide.

undertaken. After fine-tuning the survey instrument, the main survey yielded 252 usable replies. Post-hoc checks (see below) showed that the data collected were highly reliable. Our experience in India is perhaps not very different from what we observed in other emerging markets.

Following the best practice in comparative research, we next assessed the measurement invariance of our SI in the above five emerging markets.

## Methodology

### *Sample data and participants*

Data analysed in this section came from SIs of the five emerging markets as described above. Each questionnaire used in each of the emerging markets is unique in the sense that some sections and their respective questions have been developed according to the idiosyncrasies of the local market. However, all five SIs had common sections and questions, which allowed us to collate directly comparable data across all five countries. Although data coming from some of the aforementioned SIs have been used in previously published work, this comparative study is the first to utilize the common parts across all five questionnaires combining the survey responses. Another unique feature of this study is the fact that, as opposed to the vast majority of extant studies which examine cross-country invariance, our study does so by drawing on a diverse set of countries and time periods. Specifically, data from these studies were collected through different time periods, ranging from 2010 to 2018. The sample can thus be considered unique since it is not only applied in different markets, but also in different time periods.

### *Constructs*

A core objective of this study was to assess whether the replication of specific sections of the SIs in various country settings can produce similar validity and reliability scores among all samples, or whether significant deviations among these scores are observed. The fact that a number of sections and questions are identical across all five SIs allows us to proceed to such a comparative analysis with regard to the validity and reliability of HRM-related questions. The five SIs, each one reflecting the responses of HR directors (or similar roles) in a specific country context, share the following iden-

tical sections and questions: (1) *HR director role – Recruitment, training and retention* (qualifications, personal characteristics, training methods, promotion performance criteria, senior executive succession methods) and (2) *Appraisals, incentives and rewards* (salary differentials, reward methods for retaining key staff, social and psychological benefits, performance criteria). All questions were measured on a five-point Likert scale, ranging from 1 (not applicable) to 5 (always applicable) in case of the *Recruitment, training and retention* section and from 1 (not important) to 5 (very important) in case of the *Appraisals, incentives and rewards* section. More details on the measures of the SIs and HRM-related questions can be found in past papers (see, for example, Darwish, Singh and Wood, 2016; Darwish *et al.*, 2020; Mohamed *et al.*, 2013; Singh, Darwish, and Anderson, 2012; Singh, Mohamed and Darwish, 2013; Singh *et al.*, 2017).

### *Reliability testing*

Our study's aim was to test whether cross-country invariance could be established. However, we first had to assess the construct reliabilities and validities. In so doing, we use a confirmatory factor analysis and test the eigenvalue, the average variance extracted and the reliability score (Cronbach's alpha) for the pooled sample as well as for each of the five groups (i.e. countries). After assessing the construct reliabilities and validities, the next step was to examine cross-country invariance across different groups.

### *HRM parameters measurement invariance*

One of the key considerations in conducting any comparative research is ensuring that the data collection instruments are valid and reliable across different countries in which the data will be collected (Schmitt and Kuljanin, 2008). In other words, the researchers have to establish whether the SI measures the same constructs regardless of the sample used or the time point at which the data were collected (Somaraju, Nye and Olenick, 2022; Vandenberg and Lance, 2000). Establishing measurement invariance (or lack thereof) has important implications for theory development (Somaraju, Nye and Olenick, 2022), in our case specifically in the area of comparative HRM. If an SI measuring key HRM parameters is characterized by measurement invariance across different

countries and/or across different points in time, then the SI measures the HRM parameters in a culturally universal way and the data on such parameters can be compared. If, however, measurement invariance cannot be established, then the assessment of the HRM parameters is contingent upon contextual idiosyncrasies and future research would have to establish the underlying reasons for such differences (Somaraju, Nye and Olenick, 2022). Testing of measurement invariance is of critical importance in cross-cultural research in general and HRM cross-cultural research in particular, in order to properly determine whether the psychometric properties of an SI are valid and replicable across different cultural and national groups (Gallant and Martins, 2018).

In this study we explore the measurement invariance of a particular SI that was developed to measure different HRM parameters in five emerging markets. Overall, measurement invariance for six constructs (i.e. *Personal characteristics, Qualifications, HR director role, Reward methods for retaining key staff, Training methods and Benefits and salary differentials*) is being tested. For these constructs, unidimensional models shaped the assessment of the configural model. Our decision to form unidimensional models is based on two reasons. First, according to Riordan and Vandenberg (1994), assessment of discriminant-convergent validity is considered premature at this phase, thus placing separate metric invariance analyses for each construct as more appropriate. Second, a multiple-factor solution does not seem to work in our case as no convergence is achieved due to collinearity problems between factors. Accordingly, a unidimensional factor structure is preferred. The operationalized items for each construct can be seen in the Appendix (see Table A1).

Invariance analysis tests whether the scaling and representativeness of constructs differ across groups (Jöreskog, 1971). In our study, we estimate three constrained models against an unconstrained (configural) model across all groups. This is a widely accepted method in assessing measurement invariance and requires that equality (invariance) of factor loadings, intercepts and residuals is enforced (Steenkamp and Baumgartner, 1998). Specifically, the process requires the assessment of a configural model as a baseline model that is tested against metric, scalar and residual models for each of the six examined constructs. The configural model is the fully unconstrained

model where all loadings, intercepts and residual variances are freely estimated. The metric model constrains loadings to be equal across groups, while intercepts and residual variances are freely estimated. The scalar invariance model constrains both loadings and intercepts to be equal across groups, while residual variances are freely estimated. Finally, the residual model constrains loadings, intercepts and residual variances to be equal across groups (Theoharakis and Hooley, 2008). To assess the presence of invariance across groups, we examine the effect of the equality constraint. In so doing, we compare the chi-square statistic ( $\chi^2$ ) for each constrained model. Measurement invariance is established when an insignificant change in chi-square statistic ( $\Delta\chi^2$ ) is demonstrated. Further to the change in chi-square statistic, we also assess the following fit indices: CFI (comparative fit index), TLI (Tucker–Lewis index) and RMSEA (root mean squared error of approximation). The three fit indices should be within commonly acceptable levels (i.e.  $TLI \geq 0.90$ ,  $CFI \geq 0.90$  and  $RMSEA \leq 0.08$ ) (Vandenberg and Lance, 2000).

## Results

### *Reliability analysis*

To assess the pattern of responses to each question, and specifically the items comprising a construct, we proceeded with the estimation of the principal components related to the items of each construct, specifically looking at the eigenvalue, the reliability score of the construct (Cronbach's alpha) and finally the average variance explained (AVE). Table 1 presents the average reliability and validity scores (i.e. eigenvalue, AVE and Cronbach's alpha) for each of the five countries examined, as well as for the pooled sample (i.e. responses from all five countries combined). Table 2 presents the descriptive statistics and correlation coefficients among constructs. The statistics show that the vast majority of constructs produce relatively reliable scores. First, the eigenvalue of the extracted single factor is, for most of the constructs, fairly high. Second, the reliability score<sup>5</sup> (Cronbach's alpha) is predominantly higher than or close to 0.7. The

<sup>5</sup>We also used additional metrics for measuring the reliability score of our constructs, such as the construct reliability and the Kaiser–Meyer–Olkin (KMO) test. The aforementioned metrics returned very similar scores.



Table 1. Construct validity statistics applied to five subgroups and pooled sample

	Jordan	India	Brunei	UAE	KSA	Pooled sample
<i>Personal characteristics</i>						
Eigenvalue	3.726	2.905	4.065	3.106	1.944	3.964
AVE	0.621	0.484	0.678	0.518	0.277	0.661
Alpha ( $\alpha$ )	0.866	0.778	0.889	0.785	0.480	0.894
<i>Qualifications</i>						
Eigenvalue	3.696	2.747	3.075	2.200	3.280	3.709
AVE	0.436	0.458	0.365	0.308	0.547	0.618
Alpha ( $\alpha$ )	0.872	0.746	0.806	0.633	0.829	0.872
<i>HR Director role</i>						
Eigenvalue	3.126	3.776	3.503	3.540	2.686	4.440
AVE	0.343	0.539	0.337	0.383	0.384	0.634
Alpha ( $\alpha$ )	0.769	0.857	0.820	0.831	0.874	0.902
<i>Reward methods for retaining key staff</i>						
Eigenvalue	2.894	3.463	2.850	3.474	2.297	2.867
AVE	0.349	0.577	0.344	0.579	0.377	0.478
Alpha ( $\alpha$ )	0.779	0.852	0.777	0.850	0.605	0.776
<i>Training methods</i>						
Eigenvalue	1.959	3.025	1.939	1.987	2.874	2.712
AVE	0.392	0.605	0.388	0.397	0.575	0.542
Alpha ( $\alpha$ )	0.602	0.828	0.531	0.585	0.785	0.781
<i>Benefits and salary differentials</i>						
Eigenvalue	3.837	4.035	3.548	3.717	2.279	3.870
AVE	0.479	0.504	0.443	0.464	0.285	0.483
Alpha ( $\alpha$ )	0.834	0.856	0.760	0.829	0.525	0.841

Table 2. Descriptive statistics,  $\alpha$  coefficients and Pearson's correlations

	Mean	Std dev.	Alpha ( $\alpha$ )	1	2	3	4	5
1 Personal characteristics	3.781	0.919	0.894					
2 Qualifications	3.820	0.938	0.872	0.805				
3 HR director role	3.695	0.909	0.902	0.727	0.831			
4 Reward methods for retaining key staff	3.824	0.757	0.776	0.510	0.524	0.519		
5 Training methods	3.513	0.938	0.781	0.639	0.761	0.760	0.555	
6 Benefits and salary differentials	3.730	0.809	0.841	0.765	0.767	0.735	0.651	0.701

Note: All correlations are significant at  $p < 0.001$ .

rule of thumb is that the generally accepted lower limit is 0.7, though a value of 0.6 can be deemed acceptable in exploratory research and in case of having a small number of items in the scale (Hair *et al.*, 1998). Third, although the rule of thumb suggests that AVE should ideally be greater than 0.5, previous studies suggest that values lower than 0.5 can be accepted, as long as the reliability score is greater than 0.6. In that case, the convergent validity of the construct is considered sufficient (Fornell and Larcker, 1981).

Regarding our analysis, while the reliability scores of the pooled sample were all well above the threshold of 0.7, a few country-specific con-

structs fell below the threshold of 0.6. As far as the Brunei sample is concerned, the construct on *Training methods* produced a Cronbach's alpha equalling 0.531. The same pattern was observed for the *Benefits and salary differentials* construct of the KSA sample, where the Cronbach's alpha equals 0.525. Despite these two cases, fit indices for the rest of the constructs were within acceptable levels.

Overall, our comparative analyses on reliability and validity showed that, with the exception of a few cases, the vast majority of the items used under each question seem to produce acceptable reliability scores across all samples, thus indicating

Table 3. Measurement invariance tests across countries

Construct	Model	$\chi^2$	df	$\Delta\chi^2$	$\Delta$ df	p-Value	CFI	TLI	RMSEA	Equality supported
Personal characteristics	Configural	2033.69	92				0.997	0.990	0.045	–
	Metric	2220.75	97	187.06	5	0.000	0.940	0.900	0.142	No
	Scalar	2595.83	102	375.09	5	0.000	0.777	0.761	0.220	No
	Residual	2688.49	107	92.66	5	0.000	0.767	0.816	0.193	No
Qualifications	Configural	1313.14	65				0.998	0.983	0.061	–
	Metric	1333.32	69	20.18	4	0.005	0.987	0.975	0.075	No
	Scalar	1429.50	73	96.18	4	0.000	0.944	0.938	0.118	No
	Residual	1455.90	77	26.40	4	0.000	0.936	0.951	0.109	No
HR director role	Configural	1471.83	124				0.975	0.933	0.079	–
	Metric	1476.74	130	4.91	6	0.555	0.976	0.965	0.072	Yes
	Scalar	1583.81	136	107.07	6	0.000	0.948	0.945	0.090	No
	Residual	1652.78	142	68.97	6	0.000	0.928	0.942	0.094	No
Reward methods for retaining key staff	Configural	453.97	68				0.979	0.923	0.074	–
	Metric	472.00	73	18.03	5	0.002	0.931	0.885	0.090	No
	Scalar	486.33	78	14.33	5	0.013	0.914	0.908	0.080	No
	Residual	1060.81	83	574.47	5	0.000	0.880	0.906	0.081	No
Training methods	Configural	1391.82	65				0.998	0.979	0.051	–
	Metric	1480.22	69	88.40	4	0.000	0.921	0.842	0.142	No
	Scalar	1624.45	73	144.23	4	0.000	0.759	0.732	0.185	No
	Residual	1720.32	77	95.86	4	0.000	0.680	0.754	0.177	No
Benefits and salary differentials	Configural	1318.73	109				0.975	0.925	0.078	–
	Metric	1449.27	116	130.54	7	0.000	0.856	0.749	0.142	No
	Scalar	1648.71	123	199.44	7	0.000	0.731	0.673	0.161	No
	Residual	2665.92	130	1017.20	7	0.000	0.634	0.659	0.165	No

Note: CFI = comparative fit index; TLI = Tucker–Lewis index; RMSEA = root mean squared error of approximation. Configural model assumes loadings, intercepts and residual variances to be freely estimated. Metric model constrains loadings to be equal across groups. Scalar invariance model constrains both loadings and intercepts to be equal across groups. Residual model constrains loadings, intercepts and residual variances to be equal across groups.

conformity between constructs across different country contexts.

#### Measurement invariance: Comparing constructs across sample groups

The process of testing measurement invariance starts by estimating unconstrained (baseline) models for each construct where no cross-group constraints have been imposed. Table 3 presents measurement invariance tests across country groups for all six constructs. Fit indices of the unconstrained models for all six constructs are within acceptable levels (i.e.  $TLI \geq 0.90$ ,  $CFI \geq 0.90$  and  $RMSEA \leq 0.08$ ). Once it was demonstrated that the fit indices of the unconstrained (i.e. configural) models were accepted, we proceeded to the estima-

tion of the constrained models (metric, scalar and residual). Constrained models assume equality (invariance) of factor loadings, intercepts and residuals. For each construct, we assessed measurement invariance across the five sampled countries.

The assessment of the constrained models did provide some interesting results. First, only three of the constrained models produced fit indices which can be deemed to be within acceptable levels. These are (1) the metric model of the *Qualifications* construct; (2) the metric model of the *HR director role* construct; and (3) the scalar model of the *Reward methods for retaining key staff* construct. For the rest of the constrained models, while both CFI and TLI statistics were within acceptable levels for most of the constrained models, RMSEA was not, thus signalling a problematic model fit overall. Second, even if the model fit was in an accept-

able range, we failed to observe a non-significant change in chi-square statistic ( $\Delta\chi^2$ ). Sequential tests constantly produced a statistically significant decrement in fit, which indicates that measurement invariance could not be established for any of the constructs. The only exception where the measurement invariance could be claimed was the metric model of the *HR director role* construct, which was improved compared to the configural model of the construct and the change in chi-square statistic was non-significant. Overall, the findings suggest that survey participants originating from (a) different contexts and (b) diverse time periods were likely to draw on vastly different conceptual structures and notions when asked to assess several HRM-related features. Therefore, invariance could not be fully established for any of the six constructs across five countries, despite the fact that some of these share certain contextual similarities.

### *Sensitivity analysis*

Given the idiosyncrasy of our research setting (i.e. emerging markets) and data collection process (i.e. data have been collected between 2010 and 2018), one might wonder whether the lack of invariance was due to the presence of variability in the data collection process (e.g. it could be the case that in one country the sample consisted of more small and medium-sized enterprises with less developed HRM practices while in another country the sample focused more on larger companies with more advanced HRM practices).<sup>6</sup> To address this issue, we first analysed the descriptive statistics of two key variables: firm size and internationalization. We focus on these two variables as growing firms and internationalizing firms are more likely to adopt advanced HRM practices (Khavul, Benson and Datta, 2010). Both variables are in dichotomous formation. Specifically, firm size takes the value 1 if the firm has more than 500 employees and the value 0 otherwise; internationalization takes the value 1 if the firm has international activity and the value 0 if it only has domestic operations. The descriptive statistics show that firm size across countries was rather unbalanced. Specifically, while large firms were sufficiently well represented in the samples of the UAE (36.2%), KSA (49%) and India (54.7%), the samples of Brunei

(3.9%) and Jordan (8.1%) were significantly under-represented by large firms. In terms of internationalization, the total sample was comparatively more balanced as the international activity spanned from 31.3% (KSA) to 57.9% (India), while the only sample which was considerably under-represented was the one from Jordan (9.1%). To alleviate any concerns related to the differences observed across countries in relation to the firm size and internationalization, we proceeded to a sensitivity test by removing the samples of Jordan and Brunei from the measurement invariance testing. The revised analysis did not substantially change the results compared to the previous analysis of measurement invariance, because none of the fully constrained models produced fit indices within acceptable levels. We also failed to observe any significant change in the chi-square statistic ( $\Delta\chi^2$ ). Therefore, our measurement invariance estimates are consistent.

## **Discussion and conclusions**

The main aim of this study was to critically discuss, analyse and review the benefits and challenges of using surveys in studying HRM parameters across different emerging markets and to provide recommendations for HRM survey design in these national contexts. Our recommendations below are not drawn from the literature or anecdotal experience alone, but are underpinned by a rigorous test of measurement invariance of an HRM instrument that was used in five different emerging markets. In so doing, we contribute unique findings to the HRM divergence–convergence debate using a relatively homogenous sample and provide recommendations for survey design that can drive future comparative HRM research.

### *Key challenges in replicating HR SIs across emerging markets: What did we learn?*

Based on our findings, we can derive two core implications. First, the lack of measurement invariance in our SI to measure the HRM parameters means that the data in each country should have been collected with a unique SI developed specifically for the country under consideration (Schmitt and Kuljanin, 2008; Somaraju, Nye and Olenick, 2022). In other words, in order to provide robust comparative findings related to different HRM parameters across the five studied emerging

<sup>6</sup>We would like to thank one of the reviewers for making this suggestion.

markets, the SI should have been developed paying attention to specific national, institutional features rather than using the same survey in all countries. This finding implies that (a) HRM practices applied in one country and/or (b) the method of approaching firms/respondents to take part in the survey are not entirely applicable in different contexts (Haak-Saheem and Darwish, 2021; Rowley, Quang and Warner, 2007) regardless of how – seemingly – homogenous the sample might be. For example, a key item in the *Rewards methods for retaining key staff* construct is the potential for the employee to achieve better career prospects and development opportunities compared to firms in the same industry. This question can be particularly relevant to the context of India or the UAE, where labour mobility is more frequently observed, compared to, for example, Brunei Darussalam, where this might not be the case. Thus, adaptation of some SI questions might be required to better align with the country context. Further, in terms of approaching respondents to take part in the survey, the effectiveness of recruiting participants can differ significantly across countries based on, for example, the propensity of firms and managers in openly sharing views with researchers or other third parties, also pertaining to the idiosyncrasies of the national culture. Again, this would require that the method of approaching respondents differs depending on the cultural setting (e.g. consider using personal connections and contacts in countries where informal networks play a more important role than formal networks do).

Therefore, despite the widely established consensus among HRM scholars that emerging markets are largely similar, thus advocating for directional convergence (Budhwar, Varma and Patel, 2016), our findings do not confirm this view. This could be due to multiple factors which are unique to each national context. For instance, distinct institutional traditions in some countries like Jordan (e.g. centring on community, cultural traditions, local values and conventions) contribute significantly to shaping HRM practices (e.g. Darwish, Singh and Wood, 2016), whilst relatively weak or underdeveloped formal institutions in India could be the main reason why ‘substitutive’ informal institutions were developed instead; the latter would serve as an alternative basis for regulation of labour and potentially impact on the development of HRM practices in the country (e.g. Banerjee and Lyer, 2005; Darwish *et al.*, 2020). In ad-

dition to these distinct institutional arrangements of each of the five national contexts, there are further reasons why the HRM in the UAE, KSA and Brunei might be different. These nations are underpinned by petroleum growth regimes, with institutional foundations being centred around the promotion of oil and gas and their associated significant revenues; as suggested by resource curse theory, non-oil and gas sectors in petrostates could suffer a drain in human resources, crowding out investment in other areas (Wood *et al.*, 2020) and hence significantly affecting HRM systems and the development of human capital within these countries. Therefore, all these factors may well explain the lack of SI measurement invariance used to measure HRM parameters in five different emerging markets.

Also, in relation to the first implication, one has to consider the longitudinal setting of our research methodology. Specifically, while the vast majority of research studying cross-sectional measurement invariance of HRM parameters is conducted in a single period (i.e. within a few weeks or months), our research is not, as the data used were collected over an extensive time period. This constitutes a methodological idiosyncrasy as it adds an extra layer of complexity in the applied research methodology. This is because we have to consider not only the contextual differences across markets, but also the temporal dimension affecting the adoption of HRM practices. This calls for more emphasis being put on effectively dealing with potential temporal effects on various HRM practices in a cross-country setting. But why is this the case? Firstly, deploying an SI in different countries demands that the SI has been sufficiently adapted to the host country’s context, possibly requiring additional effort being put into studying the local context, making on-site visits, running pilot surveys in each country, partnering with different local actors, among others. This process requires time and dedicated resources being put into each country setting. Secondly, applying a seemingly replicable SI in a cross-sectional setting within a short term can result in temporal errors. This is because findings in such a cross-sectional design might not be applicable over longer observation periods, as well as because short-term studies may underestimate the strength of long-term effects (Aguinis and Bakker, 2021). Therefore, researchers should ensure that any temporal characteristics related to an SI are being set early in the design phase.

Regarding our second implication, it is crucial for studies proceeding to convergence of responses coming from different (country) samples to systematically assess the validity and reliability of constructs for each sample or subsample before doing so for the merged dataset. Yet, testing for the validity and reliability of constructs is not sufficient to confirm the potential convergence of responses. For example, in our study, although constructs have turned out to be highly reliable for the merged sample, convergence was not established. In fact, the assessment of measurement invariance clearly showed that equivalence among the different samples could not be achieved. Based on this, we suggest that, at the very least, researchers should treat each country's SIs as a standalone dataset, thus proceeding to the assessment of the validity and reliability of constructs both before and after the convergence of the merged dataset.

#### *Recommendations for future research*

Our empirical results could lead the way to a number of opportunities to develop SIs on HRM parameters and test hypotheses related to HRM and international HRM across emerging markets. To this end, our key take aways are as follows. First, the lack of measurement invariance or comparability of our findings across five emerging markets suggests that more attention should be paid to SI development for comparative research purposes. For instance, researchers could involve the HR directors from different emerging countries earlier on in the process of SI development to get their views around their companies' HRM parameters and how these could best be measured. Alternatively, researchers could focus on testing the dynamics of the HR director's role in different countries, which could explain why HRM parameters cannot be compared across different national contexts.

Second, researchers should ascertain the key recruitment criteria followed by firms, their retention practices for key staff, the modus operandi of their training methods, performance appraisal approaches, methods of executive succession, incentives and rewards methods and cross-cultural issues. These exercises could be carried out inter-country between domestic firms and MNEs (e.g. intra-country), and could potentially feed into one overall, valid and reliable SI that could be used in a comparative research setting.

Third, special attention should be paid to the content and wording of the questions, as well as choosing the right rating or response scales as these are important criteria in the initial stages to ensure the validity and reliability of the collected data (Robinson, 2018). Our experience tells us that constructing an efficient survey that allows researchers to collect valid and reliable data in different cultural contexts can be a very difficult task. Although researchers and HRM practitioners can gain a lot of information from using surveys, often poor survey design leads to inaccurate or incomplete data collection (Morrel-Samuels, 2002). An example of this would be different job satisfaction surveys which tend to show relatively high levels of satisfaction across the workforce and yet – when digging deeper into workers' experience of satisfaction by means of qualitative interviews – one finds that workers are actually not as satisfied as the surveys might suggest.

Finally, our findings on the lack of measurement invariance across the five emerging markets further confirm that best-practice HRM archetypes are more of a myth than a reality (see Mayrhofer, Brewster and Farndale, 2018). Hence, the next generation of empirical work on international and comparative HRM should develop novel context-dependent theory (see Farndale *et al.*, 2023). Based on an in-depth contextual examination, scholars need to highlight and nuance the central importance of context to international and comparative HRM research, and further interrogate existing theoretical conversations, with the aim of extending or synthesizing theory.

In addition to the above key take aways, Table 4 offers further specific recommendations for survey design that can drive future comparative HRM research in these emerging markets under study and other comparable contexts.

#### *Limitations*

One of the major limitations of the study could be the lack of proficiency in English by some respondents. This is a factor of critical importance in the process of collecting reliable, unbiased responses in an SI. Also, although the longitudinal setting of our study was a strength that allowed us to test measurement invariance across multiple time frames, it could be that HRM practices may have changed throughout time (i.e. between 2010 and 2018) in some of these contexts, thus calling

Table 4. Survey design recommendations for future comparative HRM research in emerging markets

Problem	Recommendation
<i>General problems associated with conducting HRM SI</i>	
How to stimulate research interest among local organizations?	Establish a research team of collaborators from the local area Invest in personal connections and contacts (e.g. <i>guanxi</i> , <i>wasta</i> , <i>blat</i> , etc.); where possible (e.g. small sample size or limited geographical location), following a face-to-face approach seems to be more effective due to cultural reasons
How to increase survey response rate?	
How to speed up the process of data collection on representative samples?	Consider working with a local market research collaborator or university research partner
<i>Problems associated with replicating HRM SI</i>	
How to integrate/replicate HRM SIs across different markets?	Design HRM SI considering the idiosyncrasies of each country under investigation
How to adapt an HRM SI across different markets also responding to local market idiosyncrasies?	Involve HR directors from different emerging countries earlier on in the process of SI development
How to achieve equivalence in term/word meaning across surveys in different languages?	Use parallel translation approach rather than back translation
How to use HRM SIs across countries that share profound similarities (e.g. employment law, industrial relations, etc.)?	Consider more nuanced and profound socio-political characteristics embedded in the society rather than focusing on geographical or economic indicators alone
How does time affect the design and deployment of the SI across countries?	Ensure that any temporal characteristics related to the design and deployment of the SI are being set early in the design phase
How to treat constructs resulting from different country samples in relation to a pooled (aggregate) sample?	Assess validity and reliability of constructs both before and after the convergence of the merged dataset

for more systematic ways of factoring the longitudinal idiosyncrasy into the research methodology in the future. Finally, it would be useful if future research could include more nations and regions when assessing measurement invariance. In turn, this would help in providing unique empirical findings of comparability of SIs assessing different HRM parameters across national contexts.

## References

- Aguinis, H. and R. M. Bakker (2021). 'Time is of the essence: improving the conceptualization and measurement of time', *Human Resource Management Review*, **31**, art. 100763.
- Akinci, C. and M. Saunders (2015). 'Using questionnaire surveys to gather data for within organisation HRD research'. In M. Saunders and P. Tosey (eds), *Handbook of Research Methods on HRD*, pp. 217–230. Cheltenham: Edward Elgar.
- Banerjee, A. and L. Lyer (2005). 'History, institutions & economic performance: the legacy of colonial land tenure systems in India', *The American Economic Review*, **95**, pp. 1190–1213.
- Boon, C., D. N. Den Hartog and D. P. Lepak (2019). 'A systematic review of human resource management systems and their measurement', *Journal of Management*, **45**, pp. 2498–2537.
- Brewster, C., W. Mayrhofer and F. L. Cooke (2015). 'Convergence, divergence and diffusion of HRM in emerging markets'. In F. Horwitz and P. S. Budhwar (eds), *Handbook of Human Resource Management in Emerging Markets*, pp. 451–469. Cheltenham: Edward Elgar.
- Brewster, C., W. Mayrhofer and A. Reichel (2011). 'Riding the tiger? Going along with Cranet for two decades—a relational perspective', *Human Resource Management Review*, **21**, pp. 5–15.
- Budhwar, P. S., A. Varma and C. Patel (2016). 'Convergence–divergence of HRM in the Asia-Pacific: context-specific analysis and future research agenda', *Human Resource Management Review*, **26**, pp. 311–326.
- Collinson, S. C. and A. M. Pettigrew (2009). 'Comparative international business research methods'. In A. M. Rugman (ed.), *The Oxford Handbook of International Business*, 2nd edn, pp. 765–797. Oxford: Oxford University Press.
- Cooke, F. L., A. Veen and G. Wood (2017). 'What do we know about cross-country comparative studies in HRM? A critical review of literature in the period of 2000–2014', *The International Journal of Human Resource Management*, **28**, pp. 196–233.
- Cooke, F. L., G. Wood, M. Wang and A. Veen (2019). 'How far has international HRM travelled? A systematic review of literature on multinational corporations (2000–2014)', *Human Resource Management Review*, **29**, pp. 59–75.
- Cooke, F. L., M. Xiao and Y. Chen (2020). 'Still in search of strategic human resource management? A review and suggestions for future research with China as an example', *Human Resource Management*, **60**, pp. 89–118.
- Darwish, T. K., S. Singh and G. Wood (2016). 'The impact of human resource practices on actual and perceived organizational performance in a Middle Eastern emerging market', *Human Resource Management*, **55**, pp. 261–281.
- Darwish, T. K., G. Wood, S. Singh and R. Singh (2020). 'Human resource management in India: performance and complementarity', *European Management Review*, **17**, pp. 373–389.
- Datta, S., P. Budhwar, U. A. Agarwal and S. Bhargava (2023). 'Impact of HRM practices on innovative behaviour: mediating role of talent development climate in Indian firms', *The*

- International Journal of Human Resource Management*, **34**, pp. 1071–1096.
- Dibben, P., C. Brewster, M. Brookes, R. Cunha, E. Webster and G. Wood (2017). 'Institutional legacies and HRM: similarities and differences in HRM practices in Portugal and Mozambique', *The International Journal of Human Resource Management*, **28**, pp. 2519–2537.
- Dillman, D. A. (1978). *Mail and Telephone Survey, the Total Design Method*. New York: Wiley-Interscience.
- Douglas, S. P. and C. S. Craig (2007). 'Collaborative and iterative translation: an alternative approach to back translation', *Journal of International Marketing*, **15**, pp. 30–43.
- Farndale, E., J. Bonache, A. McDonnell and B. Kwon (2023). 'Positioning context front and center in international human resource management research', *Human Resource Management Journal*, **33**, pp. 1–16.
- Fornell, C. and D. F. Larcker (1981). 'Evaluating structural equation models with unobservable variables and measurement error', *Journal of Marketing Research*, **18**, pp. 39–50.
- Gallant, W. H. and N. Martins (2018). 'Measurement invariance of employee engagement across race groups', *International Journal of Productivity and Performance Management*, **67**, pp. 1463–1481.
- Gao, Z., C. Zhao, F. L. Cooke, B. Zhang and R. Xie (2020). 'Only time can tell: whether and when the improvement in career development opportunities alleviates knowledge workers' emotional exhaustion in the Chinese context', *British Journal of Management*, **31**, pp. 206–220.
- Ge, J. and W. Zhao (2020). 'Culture, power, and hybrid HR system in foreign-invested enterprises: evidence from China', *The International Journal of Human Resource Management*, **31**, pp. 2372–2405.
- Groves, R. M. (1990). 'Theories and methods of telephone surveys', *Annual Review of Sociology*, **16**, pp. 221–240.
- Haak-Saheem, W. and T. K. Darwish (2021). 'Human resource management in the Middle East'. In E. Parry, M. J. Morley and C. Brewster (eds), *The Oxford Handbook of Contextual Approaches to Human Resource Management*, 1st edn, pp. 327–344. New York: Oxford University Press.
- Haak-Saheem, W., M. Festing and T. K. Darwish (2017). 'International human resource management in the Arab Gulf States – an institutional perspective', *The International Journal of Human Resource Management*, **28**, pp. 2684–2712.
- Hair, J. F., R. E. Anderson, R. L. Tatham and C. William (1998). *Multivariate Data Analysis*. Upper Saddle River, NJ: Prentice Hall International.
- Hambleton, R. K. (1993). 'Translating achievement tests for use in cross-national studies', *European Journal of Psychological Assessment*, **9**, pp. 57–68.
- Hofstede, G. (1980). *Culture's Consequences: International Differences in Work-Related Values*. Beverly Hills, CA: Sage.
- Hoorani, B. H., E. Plakoyiannaki and M. Gibbert (2023). 'Understanding time in qualitative international business research: towards four styles of temporal theorizing', *Journal of World Business*, **58**, art. 101369.
- House, R. J., P. W. Dorfman, M. Javidan, P. J. Hanges and M. F. S. De Luque (2013). *Strategic Leadership Across Cultures: GLOBE Study of CEO Leadership Behavior and Effectiveness in 24 Countries*. London: Sage.
- Jöreskog, K. G. (1971). 'Simultaneous factor analysis in several populations', *Psychometrika*, **36**, pp. 409–426.
- Kaufmann, D., A. Kraay and M. Mastruzzi (2009). 'Governance matters VIII: Aggregate and individual governance indicators 1996–2008', World Bank Policy Research Working Paper 4978.
- Khavul, S., G. S. Benson and D. K. Datta (2010). 'Is internationalization associated with investments in HRM? A study of entrepreneurial firms in emerging markets', *Human Resource Management*, **49**, pp. 693–713.
- Lakshman, C., L. Wang, A. Adhikari and G. Cheng (2022). 'Flexibility oriented HRM practices and innovation: evidence from China and India', *The International Journal of Human Resource Management*, **33**, pp. 2473–2502.
- Latukha, M. O. (2018). 'Can talent management practices be considered as a basis for sustainable competitive advantages in emerging-market firms? Evidence from Russia', *Thunderbird International Business Review*, **60**, pp. 69–87.
- Luo, Y. and R. Tung (2007). 'International expansion of emerging market enterprises: a springboard perspective', *Journal of International Business Studies*, **38**, pp. 481–498.
- Malik, A., V. Pereira, P. Budhwar, F. J. Froese, D. Minbaeva, J. Sun, et al. (2022). 'Multilevel relational influences on HRM practices: a cross-country comparative reflective review of HRM practices in Asia', *Asian Business Management*, **21**, pp. 745–779.
- Mariappanadar, S. (2020). 'Do HRM systems impose restrictions on employee quality of life? Evidence from a sustainable HRM perspective', *Journal of Business Research*, **118**, pp. 38–48.
- Mayrhofer, W., C. Brewster and E. Farndale (2018). 'Future avenues for comparative human resource management'. In C. Brewster, W. Mayrhofer and E. Farndale (eds), *Handbook of Research in Comparative Human Resource Management*, 2nd edn, pp. 633–650. Cheltenham: Edward Elgar.
- Mohamed, A. F., S. Singh, Z. Irani and T. K. Darwish (2013). 'An analysis of recruitment, training and retention practices in domestic and multinational enterprises in the country of Brunei Darussalam', *The International Journal of Human Resource Management*, **24**, pp. 2054–2081.
- Morrel-Samuels, P. (2002). 'Getting the truth into workplace surveys', *Harvard Business Review*, **80**(2), pp. 111–118.
- Riordan, C. M. and R. J. Vandenberg (1994). 'A central question in cross-cultural research: Do employees of different cultures interpret work-related measures in an equivalent manner?', *Journal of Management*, **20**, pp. 643–671.
- Robinson, M. A. (2018). 'Using multi-item psychometric scales for research and practice in human resource management', *Human Resource Management*, **57**, pp. 739–750.
- Rowley, C., T. Quang and M. Warner (2007). 'To what extent can management practices be transferred between countries? The case of human resource management in Vietnam', *Journal of World Business*, **42**, pp. 113–127.
- Schmitt, N. and G. Kuljanin (2008). 'Measurement invariance: review of practice and implications', *Human Resource Management Review*, **18**, pp. 210–222.
- Schotter, A. P., K. Meyer and G. Wood (2021). 'Organizational and comparative institutionalism in international HRM: toward an integrative research agenda', *Human Resource Management*, **60**, pp. 205–227.
- Singh, S., T. K. Darwish and N. Anderson (2012). 'Strategic intent, high-performance HRM, and the role of the HR director: an investigation into attitudes and practices in the country of

- Jordan', *The International Journal of Human Resource Management*, **23**, pp. 3027–3044.
- Singh, S., T. K. Darwish, G. Wood and A. F. Mohamed (2017). 'Institutions, complementarity, human resource management and performance in a South-East Asian petrostate: the case of Brunei', *The International Journal of Human Resource Management*, **28**, pp. 2538–2569.
- Singh, S., A. F. Mohamed and T. Darwish (2013). 'A comparative study of performance appraisals, incentives and rewards practices in domestic and multinational enterprises in the country of Brunei Darussalam', *The International Journal of Human Resource Management*, **24**, pp. 3577–3598.
- Skuza, A., A. McDonnell and H. Scullion (2015). 'Talent management in the emerging markets'. In F. M. Horwitz and P. Budwhar (eds), *Handbook of Human Resource Management in Emerging Markets*, pp. 225–243. Cheltenham: Edward Elgar.
- Soklov, D. and E. Zavyalova (2021). 'Human resource management systems and intellectual capital: is the relationship universal in knowledge intensive firms?', *International Journal of Manpower*, **42**, pp. 683–701.
- Somaraju, A. V., C. D. Nye and J. Olenick (2022). 'A review of measurement equivalence in organisational research: what's old, what's new, what's next?', *Organisational Research Methods*, **25**, pp. 741–785.
- Steenkamp, J. B. E. and H. Baumgartner (1998). 'Assessing measurement invariance in cross-national consumer research', *Journal of Consumer Research*, **25**, pp. 78–90.
- Theoharakis, V. and G. Hooley (2008). 'Customer orientation and innovativeness: Differing roles in New and Old Europe', *International Journal of Research in Marketing*, **25**, pp. 69–79.
- Vandenberg, R. J. and C. E. Lance (2000). 'A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research', *Organizational Research Methods*, **3**, pp. 4–70.
- Wood, G., T. Darwish, A. Alanezi and S. Singh (2020). 'Indigenization of staffing in MNEs: the case of Saudi Arabia', *Asia Pacific Journal of Management*, **17**, pp. 879–898.
- Yang, F., X. Huang, D. Tang, J. Yang and L. Wu (2021). 'How guanxi HRM practice relates to emotional exhaustion and job performance: the moderating role of individual pay for performance', *The International Journal of Human Resource Management*, **32**, pp. 2493–2518.
- Zahra, S. A. (2011). 'Doing research in the (new) Middle East: sailing with the wind', *Academy of Management Perspectives*, **25**, pp. 6–21.

Tamer K. Darwish is a Professor of Human Resource Management (HRM) and Head of the HRM Research Centre at the Business School, University of Gloucestershire. He is also an Academic Fellow of the Chartered Institute of Personnel and Development. His research interests lie in the areas of strategic HRM, international and comparative HRM, and organizational performance measurement. He has published in these areas in leading management and HRM journals.

Satwinder Singh – Ex-Professor of International Business (IB) and Strategy – holds an MA and PhD in Economics and teaches IB and strategy-related modules at postgraduate level. He is also an Associate Fellow at the John H. Dunning Centre for International Business, University of Reading, UK. He has published widely in the areas of IB, strategy and international human resource management. His 2016 paper 'Measuring organizational performance: A case for subjective measures' (*British Journal of Management*, **27**, pp. 214–224) was a top cited paper for that year.

Georgios Batsakis is Associate Professor of International Business at Alba Graduate Business School and Brunel University London. His research focuses on the internationalization processes of multinational enterprises. He has published in journals such as the *Journal of International Business Studies*, *Journal of World Business*, *Global Strategy Journal*, *British Journal of Management* and *Journal of Product Innovation Management*, among others. In 2022 he was included in the Poets&Quants Best 40-Under-40 Business School Professors in the world.

Kristina Potočnik is Professor and Chair of Organizational Behaviour at the University of Edinburgh Business School. She received her PhD in Psychology from the University of Valencia. Kristina's current research focuses on innovation and resilience, as well as healthy ageing at work. Her research has been published in the *Journal of Management*, *Organization Science*, *Journal of Occupational and Organizational Psychology*, *Human Relations* and *British Journal of Management*, among others.



## **Supporting Information**

Additional supporting information can be found online in the Supporting Information section at the end of the article.