EIBA 2010 Annual Meeting | December 9-11

Track 11. Contemporary issues on IB methodology

Title of paper:

Development of a measure of expatriate job performance

Author:

Leanda Lee

Leanda Lee, PhD
Assistant Professor, School of Management, Leadership and Government
University of Saint Joseph
16/3F, Rua de Londres, NAPE
Macao, SAR China
Macau: +853 6658 9869

Development of a measure of expatriate job performance

Abstract

Based upon Campbell's (1990) multi-factorial model of job performance the Expatriate Performance Scale was developed to measure the components of expatriate performance. Item creation for the scale was informed by job performance theory applicable to both expatriate and managerial job performance with further exploration of the construct undertaken through content analysis of data from semi-structured interviews with 20 expatriate employees. The scale development comprised a review of items and sorting by subject matter experts prior to a pilot study, the data from which were subjected to item analysis. The modified scale comprising 48 items was then administered to a sample of 106 Australian expatriate employees in the Special Administrative Regions of China. These data were subjected to reliability analysis, and examined for factorial structure using Principal Components Analysis with oblimin rotation. The scale development and validation process resulted in 32 items measuring an amended model of expatriate performance with six components rather than Campbell's original eight. This six factor structure is consistent with Campbell's position that not all jobs will incorporate all performance components.

This study contributes to expatriate performance theory by the development of a measure of expatriate performance which distinguishes between components (relevant to performance outcomes required by organisations) and predictors of job performance (such as cross-cultural adjustment). The development of a scale based upon solid deductive and inductive foundations also nullifies the criticism often applied to the domain that it is purely conceptual in nature. In practical terms this scale has potential to offer a multi-dimensional measure of expatriate

performance to determine to what extent various organisational support mechanisms and individual level variables predict variance in each of the six components of performance.

Keywords

International HRM; Expatriation; Cross-Cultural Research/Measurement Issues; scale development; quantitative research methods;

This paper describes the procedures undertaken to develop the measure of expatriate performance based upon the job performance model by Campbell (1990). This measure enables relationships between antecedent variables and the separate components of performance to be tested. This has implications for practice in that practitioners now have the tools available to determine which components of performance are important to organisational goals and offers them a performance criterion measure to test the relationships between various support mechanisms and performance components. The ability to test these relationships will help to determine the best utilisation of resources to achieve specific organisational goals. In the past, much of the focus in the expatriate domain has been upon measuring interim criterion variables such as cross-cultural adjustment (Thomas & Lazarova, 2005) which may or may not have positive effects upon organisational goals.

Campbell's (1990) performance model was selected as it was determined to be superior to existing models and frameworks found in the expatriate literature (Borman & Brush, 1993; Borman & Motowidlo, 1993, 1997; Caligiuri, 1997; Campbell, 1990; Sinangil & Ones, 2003). The model is comprehensive for three main reasons: (1) it incorporates the main elements of existing conceptualisations of expatriate job performance, (2) it is underpinned by theoretical propositions which explicate how to determine each component of job performance (Campbell, Ford et al., 1990; Campbell, McCloy, Oppler, & Sager, 1993), and (3) it was developed based upon sound empirical evidence (Campbell, McHenry, & Wise, 1990) rather than purely conceptual considerations. A combination of a rigorous approach to developing an expatriate performance measure together with a clear delineation between components of performance and predictors of performance such as cross-cultural adjustment is lacking in the domain (Arthur & Bennett, 1995; Caligiuri, 1997; Mol, Born, Willemsen, & Van Der Molen, 2005; Thomas &

Lazarova, 2005). Campbell's performance model helps to both define the scope and delimit the content of the expatriate performance construct and was thus chosen as the model of choice for this study.

A multi-factorial conceptualisation of performance would be beneficial to the analysis of what may influence expatriate performance as it is expected that the various antecedents will affect each performance component in different ways. Where a construct is heterogeneous in its make-up, such as job performance or intelligence, it is best to utilise a number of homogeneous sub-scales that offer a less ambiguous interpretation than a single heterogeneous test (Anastasi, 1982). The performance scale was therefore developed to reflect the underlying latent structure of expatriate performance.

Australia's reliance on success in international markets, particularly Asia, highlights the importance of employees' performance in that context. Australians are working overseas in ever increasing numbers and consist of a select group of people with high levels of education and income. The largest group of Australian expatriates is based in Hong Kong (Hugo, Rudd, & Harris, 2003). Given the importance of greater China for Australia's continued economic growth and the size of the Australian expatriate population undertaking work in the area, it is the performance of Australia's human resources within the Chinese context which is used as the sample for this study.

PERFORMANCE IN THE EXPATRIATE DOMAIN

Of the various ways to manage human resources and staff international organisations, the expatriate assignment is still the most common method utilised and has continued to rise over a period of some years (GMAC, 2003, 2008; PriceWaterhouse Coopers, 2005). Understanding how

to achieve successful utilisation of human resources across borders has driven expatriate research over the last three decades. The recent growth suggests that job performance within this context will have heightened importance.

Expatriates

The definition of an expatriate appears to be changing as is evidenced by the variety of definitions observable in the literature. An expatriate may be defined as any individual who relocates to a country other than the home country for purposes of work or study for more that six months on a temporary and voluntary basis (Aycan & Kanungo, 1997). To distinguish from self-employed individuals or students, an expatriate has been defined by Arthur and Bennett (1995) as any employee who relocates from his home base to an international location within the same or affiliated firm and sets up temporary residence in the host country. It appears that this conventional definition is no longer as apt as it once was. There is evidence available that a wider variety of assignment types is being utilised more readily (Fenwick, 2004) such as localisation of expatriates to local terms and conditions, permanent transfers, and commuter assignments (Brookfield Global Relocation Services, 2009; PriceWaterhouse Coopers, 2005). There are also a growing number of self-selected expatriates such as independent expatriates (Harrison, Shaffer, & Bhaskar-Shrinivas, 2004; Hipsher, 2008; Richardson & Mallon, 2005) and long term global managers (Suutari, 2003). With particular relevance to Australian expatriates, Hutchings found that her sample differed noticeably; “this cohort’s experience of an international posting differs markedly from the traditional expatriate experience in that increasing numbers of Australians working in China have self-selected.” (Hutchings, 2005, p.563). In his book Hipsher (2008) focused on these self-selected members of the mobile global workforce. Although these expatriates are less accessible to researchers than the traditional expatriate, Hipsher suggested that they are growing in number for more than economic reasons and that many move overseas as a

lifestyle choice. This was confirmed in a qualitative study of 30 academic expatriates (Richardson & Mallon, 2005). Taking into consideration this possible trend towards a wider range of expatriate type and calls for research to include these (Thomas, Lazarova, & Inkson, 2005), a more encompassing definition of expatriates based upon Arthur and Bennett's (1995) is utilised in this study. An expatriate is defined as any individual who relocates from his/her home base to an international location for business or work purposes and sets up temporary residence in the host country. This definition encompasses third country nationals (TCNs) as well as parent country nationals (PCNs); the difference being a function of the similarity or otherwise of the origins of the employing multinational corporation (MNC) and the expatriate (Harrison et al., 2004). The above definition has been deliberately chosen to allow the inclusion of a wide variety of expatriates employed or contracted by local and foreign organisations, and working for small to medium size enterprises with international interests.

Performance

Kraimer and Wayne (2004) offered a summary list of expatriate success constructs which have been examined in the literature. These constructs include remaining in the assignment until the end of the term, adjusting to living conditions, and performing well on the job. Variations of these constructs have been described and used in empirical studies but it is argued here that most performance constructs considered to date may not reliably measure the outcomes directly attributable to the expatriate, or may be viewed as independent variables related to other performance constructs (Porter & Tansky, 1999). Many empirical studies have defined performance as *finishing the assignment* (Caligiuri, 1997; Farid & Buda, 1998; Kraimer & Wayne, 2004; Zeira & Banai, 1985). Rather than address the definition of performance, some authors have discussed the reasons for *failure* (Forster, 1992; Hays, 1974; Tung, 1982; Zeira & Banai, 1985). Another definition offered in the literature is the lack of *adjustment to foreign*

country or culture which has been linked to the desire to terminate the assignment (Bhaskar-Shrinivas, Harrison, Shaffer, & Luk, 2005; Caligiuri, 1997; Harrison & Shaffer, 2005; Kraimer & Wayne, 2004). *Achieving expected outcomes* (Caligiuri, 1997) or *on the job effectiveness* (Guzzo, 1996) are other descriptors which would appear to be more relevant and directly important to organisations. The difficulties are evidenced by the lack of studies in the expatriate literature which give consideration to performance theory (Mol et al., 2005). Performance theory suggests that performance is multi-faceted (Campbell, 1990). By breaking performance into separate performance factors we may develop an understanding of the relationship between various inputs (such as selection) to the expatriate management process and expatriate performance outcomes. Ones and Viswesvaran (1997) suggested that this conceptualisation of performance is valuable to the definition and measurement of expatriate performance. With this in mind, a number of empirical and conceptual studies that offer ways to break performance down into components (Borman & Brush, 1993; Borman & Motowidlo, 1993; Caligiuri, 1997; Campbell, 1990; Sinangil & Ones, 2003) were reviewed in order to select a theoretical model as a foundation for exploring the relationships between various inputs into the expatriate management process (or antecedents) and expatriate performance outcomes.

Campbell's model offers a number of advantages over others. Firstly, Campbell (1990) made distinctions between the terms performance, effectiveness, and productivity which helps to delimit the construct. Secondly, Campbell suggests that performance is the outcome of three determinants: (a) knowledge about facts and things, termed declarative knowledge; (b) knowledge about how things are done and skills to do them, termed procedural knowledge and skills; and (c) motivation to act, motivation to expend effort and motivation to persist, termed motivation (Campbell et al., 1993). The existence of such criteria in the model enables a test against which potential variables and items can be viewed for content applicability. Lastly, the model separates

performance into eight categories of activities that people are expected to undertake in their jobs (see Table 1). Not all of these factors, however, are relevant across all jobs but Campbell suggested that this list is comprehensive and at the highest level can describe every occupational job (Campbell et al., 1993). As such, other multidimensional performance factor models such as the empirically derived nine dimensions of Sinangil and Ones (2003)and the 18 factors of Borman and Brush's (1993) managerial performance taxonomy are able to be absorbed within Campbell's model (Borman & Brush, 1993; Sinangil & Ones, 2003). These models either broaden our understanding of Campbell's model or assist in categorising Campbell's factors at a higher level such as through Borman and Motowidlo's (1993) two factors of performance: task and contextual performance where task performance is related to formal aspects of a job and contributes to the technical core directly and contextual performance is related to activities that contribute to the organisational, social and psychological context of the workplace.

One of the most utilised conceptualisations of expatriate performance in the literature is the empirically driven work on expatriate performance by Caligiuri (1997). Much of Caligiuri's initial conceptual propositions were based upon the models of Borman and Motowidlo (1993) and Campbell, McCloy, Oppler, and Sager (1993). Caligiuri broke performance down into four dimensions: Technical Performance, Contextual/Pro-social Performance, Contextual/ Managerial Performance, and Expatriate Specific Dimensions. Caligiuri's taxonomy is influential as it has been cited, although only shortened versions used (consistent with Borman and Motowidlo's categorisation of task and contextual performance in most cases) as the theoretical foundation for a number of empirical studies (Harrison & Shaffer, 2005; Kraimer & Wayne, 2004; Kraimer, Wayne, & Jaworski, 2001; Selmer, Lauring, & Feng, 2009; Shaffer, Harrison, Gregersen, Black, & Ferzandi, 2006).

There are however, a couple of concerns with Caligiuri's model. Firstly, Caligiuri raised a potential problem in the division of the contextual dimensions into pro-social and managerial behaviours. There is a blurring of task and contextual factors when measuring the performance of managers. Secondly, and perhaps more critically is the inclusion of expatriate specific dimensions. Caligiuri's (1997) expatriate specific dimensions, although particularly pertinent to expatriate assignments and offer insight into the added complexity of tasks required by expatriates on assignment, could, like others, be subsumed into Campbell's model. For example, replacement planning is a staffing and training activity which is managerial in nature (Campbell's factor 8). The other expatriate specific dimension of language and culture proficiency may not of itself be a performance factor but could be seen to enable some of the factors and can therefore be considered an antecedent rather than a component of performance.

The fact that Caligiuri (1997) and Sinangil and Ones (2003) have extensively borrowed from Campbell's Performance Factor Model in their own theoretical construction lends weight to its underlying strength and applicability in the definition and measurement of job performance. Lending further support for Campbell's model (1990), Scullen et al. (2003) and Conway (1996) suggested that it, amongst others, be employed to test whether there are finer grained distinctions of performance.

This paper thus proceeds with the explanation of the processes and decisions utilised in the development of the expatriate performance measure based upon Campbell's performance model (1990). The procedure described in this paper encompasses item generation, an initial subject matter expert survey conducted to explore the content adequacy of the items, the undertaking of interviews to explore the content domain which then contributed to further item generation, the main pilot study of the initial set of items, and an administration of the measure. This was then

followed by factor analyses and finally, a determination of the final set of items based upon factor loadings and interpretation of the meaningful content of the factors.

ITEM GENERATION AND PRE-TESTING

Sample

The sample for both the interviews and the pilot study was drawn from communications with Australian human resource managers who were included on a list of approximately 30 Human Resource Committee members offered to the researcher by the Australian Chamber of Commerce in Macau and Hong Kong. Further contact lists were compiled by the researcher from personal communication with Australians in Macau and Hong Kong, the daily English press, and Australian company websites on the internet. This database was then used as the initial point of contact and interviews were arranged with the contact party or through them with Australian expatriates and their supervisors.

Measures

Initial Item Generation for the Performance Scale

The expatriate performance literature was reviewed for existing scales which included a range of performance dimensions of expatriate and managerial performance (Borman & Brush, 1993; Borman & Motowidlo, 1993; Caligiuri, 1997; Campbell, 1990; Conway, 1999; Kraimer & Wayne, 2004; Kraimer et al., 2001; Scullen et al., 2003; Williams & Anderson, 1991). This review was conducted with the goal of developing a final scale with a minimum of three items for each of the eight performance components suggested by Campbell (1990). Scales comprised of as few as three items have been able to produce adequate internal reliabilities and sample the domain of interest (Hinkin, 1995; Tett, Guterman, Bleier, & Murphy, 2000). Considering the

multiplication of the number of items over the eight factors (Campbell, 1990), the risk of response fatigue and bias that long scales are said to produce (Anastasi & Urbina, 1997) was considered high should six or more items have been chosen. In selecting the items, care was taken to ensure that each performance factor was determined by declarative knowledge, procedural knowledge, and motivation as theorised by Campbell and mentioned above (1993). This became an extremely good calibrator against which to check for items that measured antecedents to performance (such as cross-cultural adjustment), rather than components of performance, which are often included in performance measures to the detriment of conceptual clarity. The items from extent scales were sorted into the eight performance factors. Where items of similar content were repeated in two or more existing scales, the item that best matched Campbell's description of that factor was selected. This process generated 26 items. Of these 26 items, six sets of three items were used to describe Campbell's factors 1, 2, 3, 6, 7, and 8 and two sets of four items described factors 4 and 5 (see Table 1 for a description Campbell's eight factors and sample items).

Table 1

Sample items from the Expatriate Performance Scale

Campbell's performance factors	Sample item
	In your work you:
1. Job specific task proficiency	fulfil the responsibilities defined in the job description.
2. Non-job specific task proficiency	show proficiency in general tasks expected of all employees at this level.
3. Written and oral communication	effectively inform leaders, team-mates and direct reports.
4. Demonstrating effort	persevere when times are tough.
5. Maintaining Personal Discipline	refrain from complaining about insignificant things.
6. Facilitating peer and team performance	encourage and foster cooperation among team members.
7. Supervision/leadership	influence subordinates to achieve goals set for them
8. Management/administration	organise and control utilisation of non-human resources

Instrument Pre-testing

The original 26 items were included in an initial pilot questionnaire which asked respondents to rate the performance of expatriates on the items and their perception of the importance of each item. This was administered to 14 expatriates and supervisors during interviews for their review of the face validity of the items and to ascertain their understanding of

the questions and their thoughts on the formatting and ease of response. It was determined from the feedback that no amendments were required to the formatting or content at this stage.

The content of the items was also reviewed by subject matter experts (SME) for clarity, appropriateness, and content validity as it has been suggested that "items that comprise a new measure should always be presented for examination" (Hinkin, 1995, p. 971). The SMEs had academic expertise in either expatriate or performance management: one senior academic, one early stage career academic, and one PhD student/HR practitioner. This composition is acceptable as sorting is a cognitive process that Hinkin deemed appropriate for those of "intellectual ability rather than work experience" (1995, p.971).

The SMEs responses raised issues of ambiguity of items and difficulty in allocating items to the factors intended by the researcher. This is entirely consistent with the existing literature on performance components where authors offer varying opinions as to which factors the underlying dimensions represent (Motowidlo & Schmit, 1999; Motowidlo & Van Scotter, 1994). Amendments were made to item wording and one item was reallocated to another component based upon the SME responses.

Although a pilot study of the initial pool of items had been commenced through administration to 14 expatriates and supervisors, it was deemed that given a lack of consensus by the SMEs, the existence of some ambiguous items, and the lack of agreement between the theory and the SME's categorisation that further work needed to be conducted. Subsequently, a second pilot test was administered after further item generation and a second SME categorisation process.

Further Item Generation

After the review by SMEs suggested that it would be prudent to build in redundancy to allow for deletion of items after subsequent factor and reliability analyses (DeVellis, 2003), an

additional 16 items were included. These were items that had been considered but not included in the original pool. Where insufficient items were generated from existing scales the author created others based upon Campbell's description of the factors. On the basis of a second review of experts, items that appeared to be ambiguous or subject to response bias were rewritten. It was also viewed as important to increase the number of items for each factor to approximately six (Tabachnick & Fidell, 2001) to build in redundancy (DeVellis, 2003). This initial screening process resulted in a pool of 42 items in the initial test instrument.

The list of items was further complemented by a content analysis of interviews with 18 Australian expatriates and two supervisors conducted in Hong Kong (11) and Macau (9) in 2006. Interviews were conducted with a major aim to discern whether the content of the performance items generated for the performance scale was relevant to the sample and to detect whether there was any content from the domain that might be missing.

The interviews were semi-structured and open-ended, organised around the content of the Expatriate Performance Framework. The questions were designed to identify the components of performance which were important to the expatriates. The first two open-ended questions aimed to ascertain the components of performance without forcing the theoretical framework on the discussion. Next, questions were asked based upon Campbell's (1990) 8 factors to test their applicability to the Special Administrative Regions (SARs) context. A combination of deductive and inductive approaches was taken in the sorting and analysis of the interview texts. Although a deductive approach was originally taken by predefining the codes in order to check the relevance of the selected variables to the population, an inductive approach was also taken to minimise the rigidity imposed by such a framework and to allow for other themes to emerge from the data (Miles & Huberman). As recommended by Dougherty and Hardy (1996), the data were sorted into themes and then the frequency of each of the themes was inspected to determine their

strength. It was clear through the coding process that there were a number of emerging patterns that were not covered by the existing items. Any text content that was not represented by existing items was then reviewed to see if the themes represented were recurring and suggestive of additional item generation. These emergent themes were coded to new nodes and then items for the scale were developed from their content. From this process another six items were added bringing the total to 48 items covering eight performance factors.

Subsequent to this content analysis of the interviews and the final addition of items and some minor adjustments to wording and formatting, the questionnaire was further reviewed by two research academics with expertise in scale and questionnaire development. After review the scale was determined to be ready for a second pilot study.

Second Pilot Study – Sample and Measure

The 48 items were pilot tested for reliability and internal scale consistency with 32 expatriates employed in a range of occupations. There were 17 (53.1%) Hong Kong based respondents, while 15 (46.9%) were from Macau. In terms of gender, 13 (40.6%) of respondents were female, while 19 (59.4%) were male.

This scale was designed to rate performance of expatriates on 48 items over eight components of performance: at this stage item redundancy was still built in. Items measuring each variable factor were mixed to overcome response set error (DeVellis, 2003).

There is a dynamic tension between specificity and generality (DeVellis, 2003) in relation to performance scales (Tett et al., 2000) which are administered across industries, functions, and roles. Although this scale samples more of the domain of interest than extent expatriate performance measures, it is still somewhat generic in order to be applicable across a wide range of

expatriate roles and to enable administration to a sample size large enough to enable generalisation..

The performance scale utilised a 5-point Likert-type response scale, ranging from 1 (poor) through to 5 (outstanding). An example of one item for each factor is listed in Table 1 above. The aim was to administer the Expatriate Performance Scale to both expatriates to obtain self-ratings and to their supervisors to obtain a more objective view (Lester, Turnley, Bloodgood, & Bolino, 2002; Scullen et al., 2003).

Second Pilot Survey Procedure

The researcher contacted the individuals who participated in the interviews, representatives of the interviewee's organisations, and other Australian expatriates to request participation in the pilot study. Of the 50 individuals contacted 35 responded, with 33 returning the survey. Due to substantial missing data, one survey was unusable, leaving 32 useable surveys for a response rate of 66%. As none of the remaining 32 cases was found to have excessive missing data, missing data were dealt with using listwise deletion. Thus most of the analysis was conducted on 31 cases.

Method of Analysis

This second pilot study was conducted for the purpose of assessing the initial psychometric properties of the survey. This was done in order to determine whether the scale and the sub-scales (grouped items aimed to measure the components of performance) were reliable enough to justify further development and administration to a larger sample.

The addition of items based upon the content analysis of the interview data was undertaken with the aim to improve the content validity. Further content validity assessment was conducted during the process of factor analysis as the interpretation of the rotated factors is dependent upon

the meaning that can be derived from the content of the items which load onto the individual factors. Other types of validity can only be determined through analysis with larger samples as conducted using the full data set. The following section thus focuses on the reliability measure and internal consistency of the performance component scales.

In order to determine the reliability of the component scales, each was examined using the reliabilities procedure (Nunnally & Bernstein, 1994). The initial analyses (see

Table 2) showed that the sub-scales met the traditional standards of reliability of a minimum Cronbach's alpha coefficient of .70 (Nunnally, 1978): $\alpha = .89$ for the composite performance ratings. The component scales (otherwise called performance dimensions, components or factors) also indicated, in general, very good reliability coefficients with room to improve with item deletion. For a newly developed measure with little manipulation of items, the scales certainly showed promise. Although it has been suggested that internal consistency coefficients less than .70 are less than adequate (Hinkin, 1995), those greater than .60 are considered acceptable for newly developed measures (Hair, Anderson, Tatham, & Black, 1998).

Performance component 5 which aimed to measure the maintenance of personal discipline appeared to be the most problematic ($\alpha = .28$ in the performance rating scale). Even this improved to $\alpha = .55$ after deletion of 2 items.

Table 2

Scale reliabilities of the performance components

Subscale	Cronbach's alpha	Highest alpha if item deleted	Number of items
Performance			
Component 1	.77	.80 (item 27)	6
Component 2	.58	.59 (item 23)	6
Component 3	.66	.71 (item 30)	7
Component 4	.69	.69 (item 4)	7
Component 5	.28	.41 (item 22)	6
Component 6	.80	.82 (item 14)	5
Component 7	.69	.75 (item 8)	5
Component 8	.71	.73 (item 47)	6

Note. N = 32

The reliability procedure conducted on a sample of 32 expatriates indicated that the scale was reliable. Further psychometric analyses were conducted on the larger sample in order to confirm the reliability of each scale. Further, both deductive and inductive approaches have been used to attempt to ensure content validity, which, along with reliability, is a minimum pre-requisite for testing validity (Hair et al., 1998; Hinkin, 1995). Although the content validity of the Expatriate Performance Scale was built in at the item generation (Anastasi & Urbina, 1997) and pre-test stages, construct validity needed to be tested through factor analyses. It was anticipated that some items would be found to be redundant or not load onto the factors as theorised and may

need to be removed from the scale. These tests are discussed in the following sections on factor analysis.

SURVEY FINAL ADMINISTRATION: RELIABILITY AND VALIDITY ANALYSIS

Sample

A purposive sampling technique (Kemper, Stringfield, & Teddlie, 2003) was utilised to access Australian expatriates and their supervisors for the survey questionnaire. Convenience sampling together with opportunistic and snowballing techniques were used to increase the coverage of the population and sample size. The Australian Chamber of Commerce and Australian Association were selected for access to expatriates through member representatives in the case of the former and directly to members in the latter. Other organisations known to have Australians working for them and individuals were then approached based upon personal contacts made during the data collation period.

Given much of the literature, a sample size of 106 appears minimally sufficient for factor analysis. Arthur and Bennett state, this is "part of the challenge of doing international management research" (Arthur & Bennett, 1995). Sample size is constrained by the practicalities of the applied situation along with the requirements of correlation estimates (Arrindell & van der Ende, 1985; Zwick & Velicer, 1986). There are various guidelines in the literature with very little agreement as to what constitutes an appropriately sized sample for factor analysis (Fabrigar, Wegener, MacCallum, & Strahan, 1999; Hogarty, Hines, Kromrey, Ferron, & R., 2005; MacCallum, Widaman, Zhang, & Hong, 1999). Three hundred cases is said to be a comfortable level (Ferguson & Cox, 1993; Tabachnick & Fidell, 2001) and particularly so when mean factor loadings on each extracted factor is low (at less than .60) (Guadagnoli & Velicer, 1988) but that smaller sample sizes should be adequate if there are several item loadings above .80. The latter

was the case found in the first rotated component matrix produced in the following factor analysis for this expatriate study, which is one indication that the profile of the data will allow for a smaller sample size. Some authors have advocated advocate “rules of thumb” with a minimum number of cases for each item in the analysis with variations from four (Hair, Anderson, Tatham, & Black, 1992) and five (Hair et al., 1998) to ten or even 20 cases per item (see Hogarty et al., 2005). However, it depends upon the methods used to decide on the number of components to retain and upon the profile of the data set as to what sample size is required. In Zwick and Velicer’s (1986) study comparing decision techniques used to retain components in principal components analysis (PCA) the authors were able to offer evidence that with a reasonably high level of saturation (.80 loadings) even a small data set of 72 cases and 36 variables could be accurately factor analysed to the correct number of components and item loadings. These findings replicated earlier ones: In a study utilising two large data sets of 491 and 1,198 participants to investigate the sample size question for factor analysis using two different personality scales, Barrett and Kline 1981 (cited in MacCallum et al., 1999) suggested that sub-samples of 48 and 112 obtain a good recovery of population loadings with case to item ratios of 3 and 1.2 respectively; Arrindell and van der Ende (1985) also had similar findings with various sample sizes down to 26. Sample sizes of 78 and above for 20 items (3.9 cases to variable) and samples sizes of 100 and above for 76 items (1.3 cases to variable) showed neither sample size nor ratio of sample size to variable had any great effect on achieving the same factorial solution as the population (cited in MacCallum et al., 1999).

Further, suitability of the sample size depends upon the stability of the outcomes of factor analysis across various extraction techniques and types of rotation employed (Tabachnick & Fidell, 2007). There is evidence that this SARs data set is relatively stable across techniques and that there exist a good number of factor loadings above .55 which is a conservative level at which

the loading can be deemed statistically significant for a sample size of 100, at a power level of 80 and .05 significance (Hair et al., 1998).

Methods of Analysis

Factor Analysis

Exploratory factor analysis was chosen to assist in the determination of construct validity and to attempt to reduce the data to the hypothesised latent variables. It is the method particularly suited to scale development (DeVellis, 2003; Hurley et al., 1997; Worthington & Whittaker, 2006). Prior to analysis the data were screened for compliance with the assumptions of normality and the existence of any outliers. Although confirmatory factor analysis requires multivariate normality ("the sum of the variables conforms to the normal curve" [Ferguson & Cox, 1993, p.86]) and all other variables in this research have been tested for this normality assumption based upon scale scores, Ferguson and Cox have stated that when conducting exploratory factor analysis it is assumed that each variable, or item, rather than the scale conforms to the normal distribution curve. It is highly unlikely, however, that with perceptual/attitudinal measures, conformity on each and every item will be achieved entirely and thus Ferguson and Cox have offered guidelines as to what is acceptable. Based on the work by Muthan and Kaplan (1985) they first suggest that the magnitude of skew and kurtosis at a level of greater than +/- 2.0 could be problematic when it occurs in more than 25% of the variables. None of the 48 variables exhibited either skew or kurtosis to that degree and thus it was determined that the data conformed to the assumption of normality for the purposes of factor analysis.

A number of outliers were found in the data utilising the *z*-score criteria in excess of ± 3.29 ($p < .001$) (Tabachnick & Fidell, 2001). All data were retained as suggested by Hair and colleagues (1998) but with amendments to the individual item scores. Cases 33, 36, 52, and 85

required amendments to a number of performance items. The *z*-scores were used to firstly check the existence of the outliers and then to calculate the new item score. As recommended by Tharenou et al. (2007) and Tabachnick and Fidell (2007) the scores were changed to one unit larger (or smaller) than the next extreme case.

The performance items were next assessed for factorability of the matrix. The correlation matrix of the 48 items was imported into an Excel spreadsheet and conditional formatting was utilised to search for correlations at both a medium level of between 0.3 and 0.5 and a large level at 0.5 and above. 52% of the correlations were above .30 which is one indication of factorability (Tabachnick & Fidell, 2001). In addition, no item failed to produce a correlation of above .30 with at least one other item being the minimum threshold required for factor analysis (Hair et al., 1998), thus all items were retained. The second assessment of factorability undertaken was Barlett's test of sphericity which at $p = .000$ passed the test of significance of $p < .05$. This test is recommended by Tabachnick and Fidell (2001) only when the case-to-item ratio is under five. As there are 2.2 cases per item in this data matrix this test is appropriate. Finally, the Kaiser-Meyer-Olkin value at .84 was above the determining level of .6 for factorability (the minimum was set at .5 by Dzuiban & Shirkey, 1974; Pallant, 2005) and at a level termed "meritorious" by Dzuiban and Shirkey (1974) in their discussion of factorability tests. Factor analysis was therefore found to be appropriate for this data set. These last three tests were conducted by undertaking a principal components analysis (PCA) with no rotation and extraction of components based on the eigenvalue greater that one rule.

Reduction of the data to determine the latent structure underlying the performance items was conducted in the first instance by looking at the output of a PCA with varimax (orthogonal) rotation. An eigenvalue of greater than one rule (otherwise known as the latent root criterion (Hair et al., 1998) or Kaiser 1 rule (Zwick & Velicer, 1986)) was selected to assist in determining how

many components to retain (DeVellis, 2003; Zwick & Velicer, 1986) and it is appropriate for data sets of between 20 and 50 items (Hair et al., 1998). Thirteen components -somewhat higher than expected- were indicated representing 74.5% of the information captured by the 48 items. Only one item had a low extraction value (communality) under .62 with the remaining items ranging from .62 to .86, suggesting that the components expressed the data well and were viewed as having high communality (MacCallum et al., 1999). A scree test (Cattell, 1966) which indicates the point at which unique variance dominates common variance suggested that two or perhaps seven components offered the best solution: two components based upon Hair and colleagues' interpretation whereby the points above the position at which the vector changes from vertical to horizontal are retained (1998); and seven based upon Nasser and associates' interpretation where the points above and to the left of a straight line drawn across the bottom of the graph are selected for retention (2002 citing Cattell and Vogelman, 1977). Kaiser's eigenvalue greater than one rule has been found in a number of studies to severely overestimate the number of components (Hayton, Allen, & Scarpello, 2004; Pallant, 2005; Zwick & Velicer, 1986) which is logical given that it was explicitly intended to indicate the upper bound of number of items to retain (Hayton et al., 2004). It was determined that the best solution would lie somewhere between two and 13 components as (a) the Kaiser rule (the upper bound) appears to be a function of the number of items (Hair et al., 1998) and tends towards overestimation and (b) reports on the inter-rater reliability on scree tests range from very good to quite poor with larger sample sizes appearing to increase the reliability (Hair et al., 1998; Zwick & Velicer, 1986).

Zwick and Velicer (1986) compared available techniques to determine how many components to retain under a variety of conditions: sample size, number of items, saturation (high or low loadings), and complexity of item loadings. In comparing the conditions in their paper with those of this performance scale data set, the results suggest that that parallel analysis may be an

accurate method of determining how many components to retain. In fact, based upon their own findings and those of other studies, Zwick and Velicer strongly discouraged the use of the Kaiser one rule (see also Fabrigar et al., 1999) and suggested that the scree test only be used in combination with other tests, contrary to the recommendations of many authors. Parallel analysis was thus conducted for this study using a statistical program developed by Watkins (2000). Parallel analysis compares the output of randomly generated eigenvalues of 100 data sets with the same number of cases and variables as the data set undergoing factor analysis, ie. n = 106 and *p* =48. The first eigenvalue from SPSS is compared with the first eigenvalue from the randomly generated parallel analysis set. If the SPSS value is greater than the eigenvalue from the parallel analysis, the component is retained. This is continued for the second, third, and subsequent values until the parallel analysis value is greater than the SPSS value at which point the component is not retained (Pallant, 2005). In comparing the eigenvalues of both unrotated PCA and unrotated axis factor analysis, two factors were recommended for retention. Given the possible existence a hierarchical structure of performance with the upper level defined as contextual and task performance (two factors) (Borman & Motowidlo, 1993) and then further divided into other performance components as suggested by Campbell's eight factors (1990), it was not surprising that a clear solution was lacking across the different methods to determine the number of factors.

To explore this solution further, the semantic content of the components was investigated as it is the theory behind the data which should be "a major guiding force" in the determination of the number of factors to retain (Nasser et al., 2002, p.399). Theoretical considerations, "multiple criteria and reasoned reflection" (Henson & Roberts, 2006, p.399), and interpretability (Worthington & Whittaker, 2006) were thus used to decide upon the number of factors to retain. The recommendation to make use of multiple methods to undertake factor analytical decisions is a recurrent theme in the literature (Hayton et al., 2004).

Orthogonally and obliquely rotated PCA and axis factor matrices extracting 2, 5, 6, 7, 8, and 10 components were exported to a spreadsheet for further analysis (the output in SPSS of 9, 12 and 13 factor loadings was also viewed but showed little meaningful groupings of items). The goal was to determine the optimal method of extraction, rotation, and number of components to retain. Each matrix was viewed for simplicity of component structure and meaningful components. Only loadings above .40 were extracted and they were sorted from the highest loading. The cut-off of .40 was the most widely used in the 60 studies examined by Henson and Roberts (2006) from four psychology journals and has been suggested by Ford and colleagues to be the commonly used rule (1986). The items loading on each extracted component were colour coded for correspondence with the hypothesised components (the initial item generation intentions). Where three or more items from the same hypothesised component loaded on the one extracted component that component was determined to represent the meaning of the hypothesised component and defined the nature of the extracted component. This was done to see to what extent the data agreed with the hypothesised factors and to ascertain the level of redundancy by determining the items that were best removed from the scale. Even though the loadings may indicate that several items were significantly related to a latent variable, "there must be a logical reason for utilizing these indicators" (O'Leary-Kelly & Vokurka, 1998, p.392) and thus reference was made back to the theoretical justification for the items. Nineteen items loaded consistently as hypothesised across the matrices extracted (known as marker variables (Tabachnick & Fidell, 2001; Zwick & Velicer, 1986)), 21 items loaded consistently but not as hypothesised and 8 items loaded weakly.

The theoretical framework suggested eight components. However, in the process of extracting each of the matrices it was noted that the items consistently loaded on six major components. Even when eight components were extracted Campbell's (1990) components

numbered two (non-job specific task proficiency) and eight (management/administration) loaded together as did components six (facilitating peer and team performance) and seven (supervision/leadership). The remaining two components were weakly determined meaning they had low loadings and poor simple structure (Hogarty et al., 2005). The mix of components six and seven across all matrices suggested that the sample of respondents viewed peer and supervisory behaviour as one component. Perhaps this indicates a view that team membership and facilitation is inclusive of peers and subordinates. Conway acknowledged that helping and cooperating with others is "not completely distinct from leadership task performance" (1999, p.7) even though he conceptualised the former as contextual performance and the latter as task performance. The loadings from the SAR data which group these items together indeed suggest that they are not viewed as distinct by the respondents. Throughout the expatriate data collection it had been difficult to operationalise non-job specific task proficiency (performance component two). Guzman and Bourke (2003) also commented upon the difficulty in differentiating between job specific and non-job specific task performance. It is not surprising therefore that the expatriates in this sample perceive general, non-job specific, tasks in a similar way to managerial and administrative tasks.

At this point one could reflect upon the results of the scree test which offered a seven component solution based upon Nasser et al's interpretation (2002). Zwick and Velicer (1986) found that scree tests consistently (over 90% of the time) over-estimated the number of factors to be retained by one or two factors. This finding suggests five or six components might be extracted in this expatriate data and is consistent with the indications from the extracted matrices that six components appear to offer a meaningful solution. Thus, a six-component solution was chosen as each of the 4 six-component matrices (varimax and oblimin rotations on PCA and axis factor analysis [FA]) offered the simplest and cleanest solutions: approximate simple structure where the

semantic content of the components made substantive sense, offering a meaningful data reduction solution. The data derived components therefore appear to reflect the hypothesised factor structure but with some adjustments which were semantically logical.

In order to decide upon the rotation to be used, the loadings in the pattern matrices (for oblimin) and the rotation component matrices (for varimax) were compared. The six factor and six component extraction matrices run under oblimin rotation offered the cleanest solutions suggesting that an oblique rotation would be the best choice. In order to help determine which rotation should be utilised, the component correlation matrix from oblimin rotation on six components was viewed for correlations between factors. There were two correlations above .3 (highest being .37) suggesting moderate correlation (see Table 3).

Table 3

Component correlation matrix

Component	1	2	3	4	5	6
1	1.00					
2	0.17	1.00				
3	0.32	0.19	1.00			
4	0.26	0.14	0.17	1.00		
5	-0.37	-0.26	-0.28	-0.28	1.00	
6	-0.29	-0.11	-0.17	-0.21	0.20	1.00

Note. Number of items: 48

Extraction Method: Principal Component Analysis.

Rotation Method: Oblimin with Kaiser Normalization.

A number of different delta settings under oblique rotation were checked for the best solution. The criteria were a clean solution which retained the maximum number of items after

deletion for low factor loadings (less than .40) and cross loadings (less than .15 difference); and conceptual consistency with other items (Worthington & Whittaker, 2006). Both PCA and FA analyses were run. The rationale for retaining items was to have enough highly reliable items with large pattern coefficients for further analysis as Velicer and Fava (1998) suggested that such retention might compensate for smaller sample sizes. It also allows selection of the most deserving items for retention.

After searching for the maximum number of items that could be retained and minimum number of cross-loadings, the choice was between PCA oblimin rotation at a delta of zero and axis factor analysis oblimin rotation at a delta of -3. After deletion of items the FA solution offered 28 items and the PCA analysis offered 32 items. The components analysis originally yielded 12 items on component number one. As suggested by O'Leary-Kelly (1998), all components were checked for uni-dimensionality by running a PCA on the items in each component. Items 15, 22, 35, and 42 loaded separately from the other items in component one. These items were also originally developed to express components other than component one and tended to have low corrected item-total correlations. These items were thus deleted. Other items in both the components matrix and factor matrix were deleted as per the criteria mentioned above.

PCA was selected as the method of choice with this expatriate data for two main reasons. Firstly, the reliability statistics obtained on the scales derived by the PCA method were slightly higher than those under the FA method. Secondly, four more items were retained in the PCA solution which offers greater breadth of component content to be captured which is important for ensuring adequate domain sampling.

After items were deleted from the PCA solution a final PCA was conducted to check the stability of the component solution (Worthington, 2006). The only change to the matrix was the

cross-loading on item 12 which under the earlier item deletion criterion of a difference of less than .15 would have been deleted. The item was retained in the second run as it had loaded moderately well. It had also loaded on the component on which it was hypothesised to load in all matrices to this point, including the FA six-factor solution.

In summary, the series of analyses and decisions indicated that a six-factor solution with oblimin rotation with delta set at zero was the "cleanest" and most interpretable. A minimum factor loading of .40 was used as a guideline for considering an item to be part of a factor. After item deletion the scales comprised 32 items. This solution accounted for 63.42% of variance. If, as Hendrick and Hendrick (1986) asserted, 44.2% of total variance accounted for by their 6 factors was healthy then the 63.4% achieved by the components in the expatriate data must certainly be sufficient. Factor pattern coefficients ranged from .45 to .92 (see table 4).

Table 4

Principal components factor analysis of 32 performance items with oblique rotation

Item No.	Items	1TP	8MA	6TL	4DE	3CP	5MD
29	Meet the job objectives	**0.80**	-0.04	-0.12	-0.01	-0.01	-0.08
7	Meet specific job responsibilities	**0.76**	0.04	-0.00	0.01	-0.10	-0.13
27	Fulfill the responsibilities defined in the job description	**0.69**	-0.08	-0.21	0.04	-0.09	-0.14
10	Meet formal performance requirements of the job	**0.62**	0.27	0.05	-0.08	-0.04	-0.25
6	Show proficiency in general tasks expected of all employees at this level	**0.60**	0.07	0.07	0.16	-0.26	0.01
34	Complete assigned duties	**0.58**	-0.10	-0.14	0.05	-0.05	-0.27
46	Consistently demonstrate effort	**0.43**	0.09	-0.07	0.26	0.01	-0.11
25	Show proficiency in in financial, quantitative and other types of analysis common to employees at his level	0.07	**0.82**	0.04	0.02	0.19	0.05
24	Organise and control utilisation of human and other resources	-0.11	**0.81**	-0.06	0.03	-0.11	-0.09
41	Organise and control utilisation of non-human resources	0.24	**0.52**	-0.23	0.05	0.07	-0.03
45	Effectively manage risk (including financial and occupational health and safety)	0.02	**0.42**	-0.33	0.08	-0.13	-0.13

(table continues)

Table 4 (continued)

Principal components factor analysis of 32 performance items with oblique rotation

Item No.	Items	1TP	8MA	6TL	4DE	3CP	5MD
1	Demonstrate technical expertise	0.38	**0.41**	-0.04	0.11	-0.23	0.22
12	Make appropriate decisions when problems occur.	0.16	**0.41**	0.07	0.14	-0.54	0.14
39	Do things to help and support his or her teammates	0.24	0.09	**-0.65**	0.05	0.07	0.04
30	Argue effectively for own position when appropriate	0.18	-0.23	**-0.64**	0.06	-0.20	0.21
37	Facilitate peer and team performance	0.12	0.14	**-0.64**	0.07	-0.03	-0.04
40	Teach and coach direct reports	-0.11	0.34	**-0.59**	-0.00	-0.18	-0.19
31	Display a personal concern for direct reports	-0.23	-0.05	**-0.53**	0.34	0.09	-0.33
9	Ensure that subordinates receive proper development	-0.18	0.37	**-0.51**	0.01	-0.27	-0.03
38	Persuade others to accept ideas	0.31	-0.02	**-0.51**	0.22	0.01	0.06
17	Influence subordinates to achieve goals set for them	-0.18	0.26	**-0.50**	0.15	-0.22	-0.18
16	Volunteer to carry out tasks that are not formally part of his/her own job	0.04	-0.04	0.04	**0.94**	-0.02	0.07
32	Volunteer to do additional tasks for the organisation	-0.04	-0.01	-0.06	**0.93**	0.06	-0.03
5	Effectively convey ideas verbally	0.04	-0.06	0.06	0.16	**-0.80**	0.01
28	Present own position clearly and decisively	0.30	-0.09	-0.28	-0.06	**-0.64**	0.04

(table continues)

Table 4 (continued)

Principal components factor analysis of 32 performance items with oblique rotation

Item No.	Items	1TP	8MA	6TL	4DE	3CP	5MD
43	Have effective communicative exchanges	0.10	-0.24	-0.15	-0.07	**-0.63**	-0.18
3	Effectively inform leaders, team-mates and direct reports	-0.01	0.29	-0.14	-0.11	**-0.48**	-0.23
11	Comply with organisational values and policies	0.06	-0.04	0.07	0.09	-0.10	**-0.74**
19	Follow orders and regulations and respect authority	0.27	0.05	0.01	-0.01	0.12	**-0.72**
8	Model appropriate behaviours	0.19	-0.05	-0.21	-0.04	-0.08	**-0.53**
20	Maintain good working relationships	0.31	0.01	0.13	0.13	-0.16	**-0.51**
4	Show dedication to the organisational goals	-0.09	0.28	0.20	0.23	-0.37	**-0.42**

Note. N=106. Oblimin rotation with delta set to zero was utilized.

1TP = Task Performance; 8MA = Managerial and Administrative Task Performance;

6TL = Teamwork and Leadership; 4DE= Demonstrating Effort;

3CP = Communication; 5MD = Maintaining Discipline

Reliability Analysis

Scale reliability tests were conducted on the 32 items comprising the six components using the above factor analytical solution. This output also offered support for the six component solution. No component was under Nunnally's (1994) alpha cut-off point of .70 (see Table 5) and all were certainly above the acceptable level of .60 for newly developed scales (DeVellis, 2003; Hinkin, 1995).

Table 5

Reliability statistics for the 6 performance sub-scales

Campbell's Components	Content of extracted components	Cronbach's Alpha	No. of items
1. Job specific task proficiency	Task performance	.90	7
2. Non-job specific task proficiency	_	_	_
3. Written and oral communication	Communication	.79	4
4. Demonstrating effort	Volunteering	.89	2
5. Maintaining Personal Discipline	Maintaining Personal Discipline	.79	5
6. Facilitating peer and team performance	Facilitating team performance (including peers and subordinates)	.87	8
7. Supervision/leadership	_	_	_
8. Management/administration	Management and Administration Performance	.81	6

Note. N=106. There were no separately extracted components for Campbell's performance components 2 and 7. These components loaded with 8 and 6 respectively.

LIMITATIONS

The scale developed was not subjected to confirmatory analysis due to the lack of a separate sample. There is a need to conduct confirmatory factor analysis on a separate and larger sample

with the same scale to confirm the construct validity of the scale and to assess the generalisability of the findings of the research.

Factor analytical techniques such as PCA utilised here are themselves tools with limitations. Factor analysis will always find factors in a data set (Babbie, 2007)} and the factors produced will not necessarily be meaningful. It then depends upon the interpretation of and decisions made by the researcher regarding factor extraction methods and criteria for retaining factors as to whether the solution creates a scale that will withstand future use under a variety of conditions. The use of this highly scientific tool is reliant upon the art, creativity, and judgement of the researcher (Cohen, 1990).

It has been suggested that redundancy must be built into scales by generating three to four times the number of items as the final scale is expected to contain (Worthington & Whittaker, 2006). With the hypothesised eight factors and the aim of including four items per factor the final scale ideally should have comprised 32 items. In this case, 128 items (32 x 4) should have been generated to allow for item redundancy. To test this number of items would likely affect response rates (Hinkin, 1995) in an already difficult environment (Hong Kong and Macau) for achieving acceptable response rates. Extant scales were used to generate the items and some scales had already been used within the expatriate domain. In addition, thorough analyses of the theory and interview data were also undertaken. Therefore, there is sufficient justification to lower the expected level of built in redundancy in this research. It may be argued that 128 items could have been reduced to a manageable number for a survey questionnaire after the SME survey or after item reduction conducted when items do not conform to assumptions of normality. However, to force production of an excessively large number of items beyond the breadth of content already generated just so they can be deleted does not offer elegance or parsimony of design. It also demands much from SMEs with already high workloads. The three or four times requirement for

item generation mentioned above (Worthington & Whittaker) is understandable within an entirely new area of research where the scale development is based upon an inductive or grounded theory approach. However, where there are existing theories and some relevant scales from which to draw domain content, the level of redundancy may be viewed as somewhat excessive (Hinkin).

The decision taken to extract six components was based upon a combination of theoretical and empirical considerations. None of the tests to determine the number of factors to extract suggested six factors directly and in effect the least desirable method, the scree test, was used and adjusted based upon the general advice that scree tests tend to over-estimate the number of components to retain. The most recommended test, parallel analysis, indicated two factors but the extracted factors did not offer a solution which was theoretically or semantically consistent. Parallel analysis can also under factor which is a limitation of the method. The impact of under factoring is more serious for smaller rather than larger sample sizes (Hayton et al., 2004). Replication of the study with a larger sample over perhaps a larger geographic area may offer consistency amongst the factor extraction tests and between those tests and the theoretical underpinnings.

SUMMARY

The general approach to the generation of items for the performance scale was a deductive one. However, an inductive approach in the form of information obtained from the interviews was used to complement the understanding of the domain and the definition as found in the literature. Additional items were added to the pilot questionnaire based upon the analysis of the interviews. The next step in the process to ensure rigour was to subject the items of the scale to sorting by subject matter experts. This process of sorting followed by replacement or deletion of items is another step towards ensuring a respectable level of content validity of the scale. Schriesheim and colleagues (1993) suggested that such processes to assess content adequacy is an immediate

necessity once items have been generated before further preparation and administration time has been invested. The next study was a pilot of the performance scale sent to the original interviewees and their colleagues. This step was undertaken to check the psychometric properties of the scale prior to the administration of the measure as part of the final questionnaire survey. The internal consistency proved to be sufficient to proceed with the items towards further development of the scale.

Once the data had been collected and tested for outliers, assumptions of normality and factorability, exploratory factor analysis was employed to test the hypothesised framework of components of performance. It was proposed that expatriate performance is multifactorial. With this sample of 106 expatriates, it appeared that the latent structure of expatriate performance was comprised of six factors.

This study contributes to expatriate performance theory by distinguishing between components and predictors of job performance. The development of a scale based upon solid deductive and inductive foundations also nullifies the criticism often applied to the domain that it is purely conceptual in nature. In practical terms this scale has potential to offer a multi-dimensional measure of expatriate performance to determine to what extent various organisational support mechanisms and individual level variables predict variance in each of the six components of performance.

References

Anastasi, A. (1982). *Psychological testing* (5th ed.). New York: Macmillan Publishing Co. Inc.

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th international ed.). Upper Saddle River, N.J.: Prentice Hall.

Arrindell, W. A., & van der Ende, J. (1985). An empirical test of the utility of the observations-to-variable ratio in factor and components analysis. *Applied Psychological Measurement, 9*, 165-178.

Arthur, W., & Bennett, W. (1995). The international assignee: the relative importance of factors perceived to contribute to success. *Personnel Psychology, 48*(1), 99-114.

Aycan, Z., & Kanungo, R. A. (1997). Current issues and future challenges in expatriate management. In Z. Aycan (Ed.), *Expatriate management: Theory and Research* (Vol. 4, pp. 245-260). Greenwich, Connecticut: JAI Press.

Babbie, E. (2007). *The practice of social research* (11th ed.). Belmont, USA: Wadsworth/Thomson Learning.

Bhaskar-Shrinivas, P., Harrison, D., Shaffer, M., & Luk, D. M. (2005). Input-base and time-based models of international adjustment: Meta-analytic evidence and theoretical extensions. *Academy of Management Journal, 48*(2), 257-281.

Borman, W. C., & Brush, D. H. (1993). More progress toward a taxonomy of managerial performance requirements. *Human Performance, 6*(1), 1-21.

Borman, W. C., & Motowidlo, S. J. (1993). Expanding the criterion domain to include elements of contextual performance. In N. Schmitt & W. C. Borman & Society for Industrial and Organizational Psychology (U.S.) (Eds.), *Personnel Selection in Organizations* (1st ed., pp. 71-98). San Francisco: Jossey-Bass.

Borman, W. C., & Motowidlo, S. J. (1997). Task performance and contextual performance: The meaning for personnel selection research. *Human Performance, 10*(2), 99-109.

Brookfield Global Relocation Services. (2009). *Global relocation trends 2009 survey report*: Brookfield Global Relocation Services.

Caligiuri, P. (1997). Assessing expatriate success: Beyond just "being there". In Z. Aycan (Ed.), *New Approaches to Employee Management. Expatriate Management: Theory and Research* (Vol. 4, pp. 117-140). Greenwich: JAI Press, Inc.

Campbell, C. H., Ford, P., Rumsey, M. G., Pulakos, E. D., Borman, W. C., Felker, D. B., De Vera, M. V., & Riegelhaupt, B. J. (1990). Development of multiple job performance measures in a representative sample of jobs. *Personnel Psychology, 43*.

Campbell, J. P. (1990). Modeling the performance prediction problem in industrial and organizational psychology. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 1, pp. 687-732). Palo Alto, CA: Consulting Psychologists Press.

Campbell, J. P., McCloy, R. A., Oppler, S. H., & Sager, C. E. (1993). A theory of performance. In N. Schmitt & W. C. Borman (Eds.), *Personnel selection in organizations* (pp. 35-70). San Francisco: Jossey-Bass.

Campbell, J. P., McHenry, J. J., & Wise, L. L. (1990). Modelling job performance in a population of jobs. *Personnel Psychology, 43*(2), 313-333.

Cohen, J. (1990). Things I have learned (so far). *American Psychologist, 45*(12), 1304-1312.

Conway, J. M. (1996). Additional construct validity evidence for the task/contextual performance distinction. *Human Performance, 9*(4), 309-329.

Conway, J. M. (1999). Distinguishing contextual performance from task performance for managerial jobs. *Journal of Applied Psychology, 84*(1), 3-13.

DeVellis, R. F. (2003). *Scale development : theory and applications* (2nd ed.). Thousand Oaks, CA: Sage Publications, Inc.

Dougherty, D., & Hardy, C. (1996). Sustained product innovation in large, mature organizations: Overcoming innovation-to-organization problems. *Academy of Management Journal, 39*(5), 1120-1153.

Dzuiban, C. D., & Shirkey, E. C. (1974). When is a correlation matrix appropriate for factor anlaysis? Some decision rules. *Psychological Bulletin, 81*(6), 358-361.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychology research. *Psychology Methods, 4*(3), 272-299.

Farid, M., & Buda, R. (1998). Enhancing the link between expatriate selection and success: Organisation commitment to global career development. *International Journal of Management, 15*(5), 516 - 523.

Fenwick, M. (2004). *On international assignment: Is expatriation the only way to go?* (Working paper ISSN 1327-5216). Melbourne: Monash University.

Ferguson, E., & Cox, T. (1993). Exploratory factor analysis: A user's guide. *International Journal of Selection and Assessment, 1*(2), 84-94.

Ford, J. K., MacCallum, R. C., & Tait, M. (1986). The application fo exploratory factor analysis in applied psychology: A crtical review and analysis. *Personnel Psychology, 39*(2), 291-315.

Forster, N. (1992). International managers and mobile families: the professional and personal dynamics of transnational career pathing and job mobility in the 1990s. *International Journal of Human Resource Management, 3*(3), 605 - 623.

GMAC. (2003). *Global Relocation Trends: 2002 Report*. Warren: GMAC Global Relocation Services.

GMAC. (2008). *Global Relocation Trends: 2007 Report*. Woodridge: GMAC Global Relocation Services.

Guadagnoli, E., & Velicer, W. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin, 103*(2), 265-275.

Guzman, E., & Burke, M. J. (2003). Development and test of an international student performance

taxonomy. *International Journal of Intercultural Relations, 27*(6), 659-681.

Guzzo, R. A. (1996). The Expatriate Employee. In C. L. Cooper & D. M. Rousseau (Eds.), *Trends in Organizational Behavior* (Vol. 3, pp. xiii, 182). Chichester, England: John Wiley & Sons.

Hair, J. F. J., Anderson, R. E., Tatham, R. L., & Black, W. C. (1992). *Multivariate data analysis with readings* (3rd ed.). New Jersey: Prentice Hall.

Hair, J. F. J., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis* (5th ed.). New Jersey: Prentice Hall.

Harrison, D. A., & Shaffer, M. A. (2005). Mapping the criterion space for expatriate success: task- and relationship-based performance, effort and adaptation. *International Journal of Human Resource Management, 16*(8), 1454-1475.

Harrison, D. A., Shaffer, M. A., & Bhaskar-Shrinivas, P. (2004). Going Places: Roads more and less traveled in research on expatriate experiences. In J. J. Martocchio (Ed.), *Research in Personnel and Human Resources Management* (Volume 23 ed., Vol. 23, pp. 199-247): JAI.

Hays, R. D. (1974). Expatriate selection: Insuring success and avoiding failure. *Journal of International Business Studies, 5*(1), 25-36.

Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods, 7* (2), 191-205.

Hendrick, C., & Hendrick, S. (1986). A theory and method of love. *Journal of Personality and Social Psychology, 50*(2), 392-402.

Hinkin, T. R. (1995). A review of scale development practices in the study of organizations. *Journal of Management, 21*(5), 967-989.

Hipsher, S. (2008). *Expatriates in Asia: Breaking free from the colonial paradigm*. Oxford: Chandos Publishing (Oxford) Limited.

Hogarty, K. Y., Hines, C. V., Kromrey, J. D., Ferron, J. M., & R., M. K. (2005). The quality of factor solutions in exploratory factor analysis: The influence of sample size, communality and overdetermination. *Educational and Psychological Measurement, 65*(2), 202-226.

Hugo, G., Rudd, D., & Harris, K. (2003). Australia's diaspora: Its size, nature and policy implications: CEDA.

Hurley, A. E., Scandura, T. A., Scriesheim, C. E., Brannick, M. T., Seers, A., Vandenberg, R. J., & Williams, L. J. (1997). Exploratory and confirmatory factor analysis: Guidelines, issues, and alternatives. *Journal of Organizational Behavior,, 18*, 667–683.

Hutchings, K. (2005). Koalas in the land of the pandas: reviewing Australian expatriates' China preparation. *International Journal of Human Resource Management, 16*(4), 553-566.

Kemper, E. A., Stringfield, S., & Teddlie, C. (2003). Mixed methods sampling strategies in social science research. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social & behavioral research* (pp. 273-296). Thousand Oaks, CA: Sage.

Kraimer, M. L., & Wayne, S. J. (2004). An examination of perceived organizational support as a

multidimensional construct in the context of an expatriate assignment. *Journal of Management, 30*(2), 209-237.

Kraimer, M. L., Wayne, S. J., & Jaworski, R. A. (2001). Sources of support and expatriate performance: The mediating role of expatriate adjustment. *Personnel Psychology, 54*(1), 71-99.

Lester, S. W., Turnley, W. H., Bloodgood, J. M., & Bolino, M. C. (2002). Not seeing eye to eye: Differences in supervisor and subordinate perceptions of and attributions for psychological contract breach. *Journal of Organizational Behaviour, 23*(1), 39-56.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size and factor analysis. *Psychological Methods, 4*(1), 84-99.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks Sage Publications.

Mol, S. T., Born, M. P., Willemsen, M. E., & Van Der Molen, H. T. (2005). Predicting expatriate job performance for selection purposes: A quantitative review. *Journal of Cross-cultural Psychology, 36*(5), 590-620.

Motowidlo, S. J., & Schmit, M. J. (1999). Performance assessment in unique jobs. In D. R. Ilgen & E. D. Pulakos (Eds.), *The changing nature of performance* (pp. 56-86). San Francisco: Jossey-Bass Publishers.

Motowidlo, S. J., & Van Scotter, J. (1994). Evidence that task performance should be distinguished from contextual performance. *Journal of Applied Psychology, 79*(4), 475-480.

Nasser, F., Benson, J., & Wisenbaker, J. (2002). The performance of regression-based variations of the visual scree for determining the number of common factors. *Educational and Psychological Measurement, 62*(3), 397-419.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

O'Leary-Kelly, S. W., & Vokurka, R. J. (1998). The empirical assessment of construct validity. *Journal of Operations Management, 16*, 387–405.

Ones, D. S., & Viswesvaran, C. (1997). Personality determinants in the prediction of aspects of expatriate job success. In Z. Aycan (Ed.), *New approaches to employee management* (Vol. 4, pp. 63-92). London, England: JAI Press Ltd.

Pallant, J. F. (2005). SPSS survival manual: A step by step guide to data analysis using SPSS for windows (version 12) (2nd ed.). Crows Nest, N.S.W.: Allen & Unwin.

Porter, G., & Tansky, J. W. (1999). Expatriate success may depend on a "Learning Orientation": Considerations for selection and training. *Human Resource Management, 38*(1), 47-60.

PriceWaterhouse Coopers. (2005). International Assignments Global Policy and Practice: Key Trends 2005. Warren: PriceWaterhouse Coopers.

Richardson, J., & Mallon, M. (2005). Career interrupted? The case of the self-directed expatriate. *Journal of World Business 40*(4), 409-420.

Schriesheim, C. A., Powers, K. J., Scandura, T. A., Gardiner, C. C., & Lankau, M. J. (1993). Improving construct measurement In management research: Comments and a quantitative approach for assessing the theoretical content adequacy of paper-and-pencil survey-type instruments. *Journal of Management, 19*(2).

Scullen, S. E., Mount, M. K., & Judge, T. A. (2003). Evidence of the construct validity of developmental ratings of managerial performance. *Journal of Applied Psychology, 88*(1), 50-67.

Selmer, J., Lauring, J., & Feng, Y. (2009). Age and expatriate job performance in Greater China. *Cross Cultural Management: An International Journal, 16*(2), 131 - 148.

Shaffer, M. A., Harrison, D. A., Gregersen, H. B., Black, J. S., & Ferzandi, L. A. (2006). You can take it with you: Individual differences and expatriate effectiveness. *Journal of Applied Psychology, 91*(1), 109-125.

Sinangil, H. K., & Ones, D. S. (2003). Gender differences in expatriate job performance. *Applied Psychology: An International Review, 52*(3), 461-475.

Suutari, V. (2003). Global managers: Career orientation, career tracks, life-style implications and career commitment. *Journal of Managerial Psychology, 18* (3), 185-207.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Boston: Pearson/Allyn & Bacon.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston: Pearson/Allyn & Bacon.

Tett, R. P., Guterman, H. A., Bleier, A., & Murphy, P. J. (2000). Development and content validation of a 'hyperdimensional' taxonomy of managerial competence. *Human Performance, 13*(3), 205–251.

Thomas, D. C., Lazarova, M., & Inkson, K. (2005). Global careers: New phenomenon or new perspectives? *Journal of World Business, 40*(4), 340-347.

Thomas, D. C., & Lazarova, M. B. (2005). Expatriate adjustment and performance: A critical review. In G. K. Stahl & I. Björkman (Eds.), *Handbook of research in international human resource management.* (pp. 247–264): Edward Elgar, Ltd.

Tung, R. (1982). Selection and training procedures of US, European and Japanese multinationals. *California Management Review, 25*(1), 57-71.

Velicer, W. F., & Fava, J. L. (1998). Effects of variable and subject sampling on factor pattern recovery. *Psychological Methods, 3*(2), 231-251.

Watkins, M. (2000). Parallel analysis: Allen and Unwin.

Williams, L. J., & Anderson, S. E. (1991). Job Satisfaction and Organizational Commitment as predictors of Organizational Citizenship Behaviors. *Journal of Management, 17*(3), 601-617.

Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist 34*(6), 806-838.

Zeira, Y., & Banai, M. (1985). Selection of expatriate managers in MNCs: The host-environment point of view. *International Studies of Management and Organization, 15*(1), 33-51.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99*(3), 432-442.