University of Bath

Development and validation of a firm-level vertical and horizontal internationalization metric

Edmund R Thompson[1,3]

Gerard H Dericks[2]

Felicia Fai[1]

May 2018

[1]University of Bath

[2]Oxford Brookes University

[3]Corresponding author: e.r.thompson@bath.ac.uk

Abstract

The lack of valid and reliable measures of firm-level vertical and horizontal internationalization is impeding the development and testing of hypothesized relationships between these respective dimensions of internationalization and a range of important MNE characteristics, actions, and effects. Through a series of qualitative and quantitative studies using data collected from senior MNE executives (total N=3,146), we develop and validate a scale to measure both vertical and horizontal firm-level internationalization. Subscales for each type of internationalization prove to be unidimensional, reliable, temporally stable, and to have predictive, cross-cultural, cross-sectoral, and discriminant validity.

Keywords: Vertical Internationalization, Horizontal Internationalization, Measurement, Scale Development.

Introduction

A venerable conceptual distinction in international business (IB) theory is that some FDI is motivated to achieve input efficiencies (vertical internationalization, VI), while other FDI is motivated to enhance sales revenue (horizontal internationalization, HI. Caves, 1971; Markusen, 1984; Helpman, 1984). Understanding the distinction between these two prime motivations of FDI remains vitally important to understanding MNEs (Benito, 2015), as underscored by VI and HI continuing to attract wide application among IB scholars (Barbopoulus, Marshall, MacInnes & McColgan, 2014; Cieślik, 2017; El-Sahli, Gullstrand & Olofsdotter 2018; Hernandez & Guillén, 2018; Jiang, Holburn & Beamish, 2018; Saini & Singhania, 2018; Sirr, Garvey & Gallagher, 2016; Stack, Ravishankar & Pentecost, 2017).

However, as Jiang et al (2018, p.10) note, IB scholars face 'a common empirical challenge' to measure VI and HI, with some scholars suggesting 'multinational firms' activities in host countries are too complex to distinguish vertical FDI from horizontal FDI' (Tanaka, 2017, p.95). Empirically operationalizing VI and HI is such a challenge that it has prompted papers with titles like 'The (im)possibility of distinguishing horizontal and vertical motivations for FDI' (Lankhuizen, 2014), and 'Hunting high and low for vertical FDI' (Davies, 2008). Heger and McCorriston (2016) find in a recent analysis of VI and HI that sometimes 'even a detailed industry classification remains uninformative' for measurement purposes (p.323). Hence, little has changed in the decade since Alfaro and Charlton (2009) observed that 'the central challenge to the literature has been the absence of a global source of firm-level data on the basis of which to distinguish between horizontal and vertical FDI' (p.2096). As Slangen and Beugeldijk (2010) usefully summarize, 'while the conceptual distinction between vertical and horizontal foreign activity is well accepted in IB, the macro-level measurement of these two main activity types has lagged behind' (p.984).

Because few countries or organizations collect and make available adequate data on VI and HI there is often a 'paucity of sufficiently disaggregated data to permit formal analyses' (Buckley, Clegg, Cross, Liu, Voss, & Zheng, 2007, p.499). For example, both the United States Bureau of Economic Analysis (BEA) and the Swedish Research Institute of Industrial Economics both collect and make available FDI data for, respectively, US and Swedish MNEs (BEA, 2017; Swedish Research Institute of Industrial Economics, 2003). But they only report aggregated data at country and industry levels to protect MNEs' potentially competitor-sensitive information, and 'data is not available for all developing and transition countries' (Sirr et al, 2016). Researchers have therefore been obliged to rely on country-specific and non-directly comparable data generally at a broad industry-level, resulting in measures of VI and HI that are often 'less precise than [scholars] would like' (Aizenman & Marion, 2004, p.133). Naturally, as Nachum and Zaheer (2005) highlight in relation to VI and HI, 'reliance on industry-level data, while it has its merits, can obscure firm-level variation' (p.758). Although some firm-level data exist, these tend to be proprietorially held and not publicly available, like the different data-sets used by Alfaro and Charlton (2009), Capolupo, Amendolagine and Serlenga (2017), and El-Sahli et al (2018). Such firm-level data are again often country-specific, usually confined to restricted industry sectors, and non-directly comparable, thereby obliging scholars to construct unique and similarly non-directly comparable proxies for VI and HI, the validities of which vary consonantly with the peculiarities of differing data-sets' details. With so many limitations in secondary datasets, it is unsurprising that some IB researchers even pose the question: 'should we use foreign direct investment data to measure the activities of multinational corporations?' (Wacker, 2016, p.980).

IB scholars faced with secondary data constraints in construct measurement usually address these directly (see Christoffersen, Plenborg, & Robson, 2014; Gerschewski & Xiao, 2015), and systematically develop and rigorously validate new measures, either based on commonly available

secondary data (e.g. Brown, Cavusgil, & Lord, 2015; Dow & Ferencikova, 2010) or that circumvent secondary data constraints via primary data collection direct from firms (see Nam, Kim, Arthurs, Sosik, & Cullen, 2016; Nes, Yelkur, & Silkoset, 2014; Pappu & Quester, 2010; Slangen & van Tudler, 2009). However, with respect to VI and HI this has not yet happened. Instead, when secondary VI and HI data has been unavailable, scholars have to date created their own measures unique to their specific research exigencies, employing differing measurement items and response formats. For example, in assessing VI in relation to FDI location determinants, Chen and Chen (1998) create a 1-item measure asking MNEs about host-country production costs using a binary categorical low/high response format. Galan, Gonzalez-Benito and Zuñiga-Vincente (2007), also investigating FDI location determinants, nevertheless assess VI differently, using a 4-item measure asking about host costs of land, labor, transportation, and raw materials, and employing not a binary categorical, but a continuous 6-point interval measure response format. Differently again, although still researching FDI location determinants, Huett, Baum, Schwens and Kabst (2014) assess VI with a 5-item measure asking about availability of universities and natural resources plus costs of energy, wages and raw materials, this time employing a continuous 5-point interval measure response format. Similar measurement variation is also apparent in direct firm-level measurement of HI (Chen & Chen, 1998; Galan et al., 2007; Huett et al., 2014). In consequence, current ad hoc VI and HI direct measurement aimed at overcoming problems with secondary data, while perhaps fit for the specific purposes to which they are put, are heterogeneous, variably valid, and unconducive to directly comparable analyses.

Theory development in IB requires valid measurement of constructs to prevent what Sullivan (1994) a quarter century ago condemned as the separation of theory from empirical hypothesis testing and phenomenon explanation. As Mesquita (2016) highlights, understanding the motives prompting firms' FDI is of prime importance to theorizing and explaining a wide range of MNEs'

activities, and few would disagree that 'differences between horizontal and vertical investments are crucial to understanding variation in the internationalization process' (Jiang et al., 2018, p.11). However, without a systematically developed and validated means directly to measure VI and HI motivations at an individual firm-level, hypothesized effects of these motivations on many aspects of MNEs will likely continue to be difficult, and sometimes impossible, to test. Indeed, with limited prospects of adequately empirically testing many such effects, they may remain theoretically unexplored and conceptually unelaborated, arguing for the objective of our paper's research: the development and systematic validation of direct VI and HI measures at the firm-level to both facilitate and stimulate more finely nuanced research in this important area.

Our research purpose, then, is to develop a valid and reliable firm-level metric of VI and HI that is parsimonious and that MNE executives can easily complete without fear of divulging specific detailed, competitor-sensitive data. We first define VI and HI construct domains to inform an appropriate measurement approach. Second, we develop and validate a schedule of items that (1) is short enough to avoid the respondent fatigue and consequent method variance sometimes associated with time-constrained senior business executive-derived data (Thompson & Phua, 2005), but long enough fully and accurately to encompass VI and HI construct domains; that (2) mitigates the effects of other sources of method variance; that (3) has content validity; (4) possesses internal consistency reliability; (5) incorporates cross-population, cross-cultural, cross-industry, and temporal stabilities; (6) exhibits criterion-related validity; and (7) that can be used both to categorize (i) firms generally and (ii) their individual country investments specifically into those constituting either primarily VI or HI, or a combination of both.

VI and HI construct definitions

Both VI and HI appear so well accepted and uncontroversial that they are often referred to with little or no attempt to define them (see Li, Scollay, & Gilbert, 2017; Kukulski & Ryan, 2011; Waldkirch, 2011). VI is frequently, and interchangeably, referred to as internationalization that is 'cost-saving' (Aronsson & Koskela, 2011, p.87), 'efficiency-seeking' (Baldwin & Okubo, 2014, p.1051), or 'resource-seeking' (Singla & George, 2013, p.2500). Whatever descriptor is used, the underlying implicit construct directly mirrors Cave's (1971) explicit conceptualization of VI being the manifestation of firms' motivations to gain and retain input efficiencies, be those inputs raw materials, components, labor, human capital, technology or other assets (Aizenman & Marion, 2004; Braconier, Norbäck, & Urban, 2005; Davies, 2008; Fukao & Wei, 2008; Di Gregorio, Musteen, & Thomas, 2008; Jabbour, 2010; Markusen & Venables, 2000; Yeaple, 2003). Hence, VI is specifically internationalization 'motivated by cost considerations' relating to the acquisition of a wide range of physical and intangible inputs (Alfaro & Charlton, 2009, p.2099).

HI is often loosely referred to by scholars as 'market-seeking', 'sales-oriented', 'buyer-accessing' FDI (Azémar & Desbordes, 2010; Barbopoulos et al., 2014; Beugelsdijk, Smeets, & Zwinkels, 2008; Kinda, 2013; Markusen, 1984; Singla & George, 2013). In such descriptors, HI is implicitly what Caves (1971) explicitly defines as internationalization motivated to exploit firm-specific assets where demand exists but cannot be exploited through licensing or exporting. Hence, HI is internationalization specifically motivated to increase revenues from sales expansion where either licensing or exporting are infeasible, ineffective or inefficient due variously to access limitations, local adaptation requirements or cost factors (Alfaro & Charlton, 2009).

Because the motivations to achieve either input efficiencies or revenue expansion via FDI underlie, respectively, VI and HI, the measurement of both constructs would seem most appropriately achieved using a reflective as opposed to formative metrical approach (Coltman,

Devinney, Midgley, & Venaik, 2008; Finn & Wang, 2014). If VI and HI were to be considered not reflectively but formatively measurable constructs, VI would in effect need to cause input efficiency motivations, and HI would need to cause revenue expansion motivations, both of which are logically implausible (Bollen & Diamantopoulos, 2017; Hardin, 2017). We hence approach VI and HI measurement reflectively rather than formatively, as have constructors of extant ad hoc direct firm-level measures of VI and HI (Chen & Chen, 1998; Galan et al., 2007; Huett et al., 2014). Accordingly, and recognizing the need to 'mount a full-scale attack on the construct validity of a new measure' (Schriesheim, Powers, Scandura, Gardiner, & Lankau, 1993, p.389), we follow standard scale development and classical test theory validation procedures (Nunnally & Bernstein, 1994).

Methods

MNE informants

Informants for all studies are senior MNE managers at both home-country headquarters and foreign operations of MNEs, their seniority identified by job titles. Country trade association and chamber of commerce membership directories in Japan, China, Hong Kong, Singapore, and Australia were used to compile a list of MNEs and their senior managers.

We define an MNE using Caves' (1996, p.1) definition if it being a firm 'that controls and manages [operations] located in at least two countries'. As Caves (2007) notes, what at the margin comprises either 'control' or 'management' or 'operations' is contentious, but such a definition implies direct investment as opposed to portfolio investment, and also implies substantially more than merely exporting or licensing through agents. This categorization is operationalized using two variables assessing whether or not a firm has (1) processing/production and/or (2) sales/marketing direct investments in one or more foreign countries.

Study 1. Item development and purging

We drafted two initial 8-item sets each reflective of, respectively, VI or HI based on a review of both discussions in seminal VI and HI literature (Alfaro & Charlton, 2009; Caves, 1971, 1996; Feenstra & Hanson, 1997; Helpman, 1984; Markusen & Maskus, 2002; Markusen & Venables, 1998, 2000; Venables, 1999) and the items used by METI in its *Survey on Overseas Business Activities* (METI, 2016). Drafted items for VI covered raw materials acquisition, securing non-raw materials resources, technology acquisition, production cost reduction, security and continuity of supplies, and re-importation to home country. Draft HI items covered sales expansion, market entry, market access, demand satisfaction, easier sales generation, and lower cost of sales generation. Items were drafted to be succinct and easily understood by senior MNE executives with a working knowledge of English. We sought to address method variance at the item drafting stage to reduce response set and second-guessing, and also added two distracter items, suggested by Podsakoff, MacKenzie and Podsakoff (2012) to help ameliorate both response set and second-guessing.

We assessed qualitatively our drafted items in discussions during interviews conducted in Japan, Hong Kong and Singapore with 19 senior executives of MNEs with home bases in Japan, Germany, Britain, Italy, the United States, Singapore, Australia and Hong Kong. This resulted in some minor redrafting and the elimination of two items interviewees suggested were largely redundant and duplicative.

To provide a quantitative basis on which to assess individual items and to purge any that performed badly, the schedule of 14 remaining substantive and the two distracter items were incorporated as a block into a questionnaire on an unrelated international business research project with senior MNE executive respondents based in Japan, Hong Kong and Singapore. Some 372 usable responses were received.

Exploratory principal component analysis was used to examine the dimensionality of the schedule and the loadings of individual items. Excluding the two distracter items and specifying a 2-component solution, the least discriminating and lowest loading VI items proved to be those relating to re-importation to home country and to security of supplies, and the worst performing HI items related to ease of sales generation and lower costs of sales generation. By eliminating these poor items, a final schedule of ten items was derived that formed two discrete components accounting for 61.02 percent of variance, with the five best performing indicators of VI having a Cronbach's alpha of .88, and the best five HI items an alpha of .80, suggesting acceptable internal consistency reliability.

The ten remaining items plus the two distracter items were entered into an exploratory principal component analysis specifying two components. This revealed a distracter item relating to competitor activities loaded least heavily on both the VI and HI components. Accordingly, for maximum brevity, only this distracter item was retained in the final schedule.

Study 2. Initial validation

Instrument, sample and procedure. In a further effort to help reduce possible method variance, items were then ordered such that VI and HI items appear alternately and the distracter item appears as item 4 so as to provide a mental brake on any potential assumption by responders that an obvious alternating pattern might exist. The resulting 11-item schedule and item order used in all subsequent validation studies, plus question-stem and 6-point agree/disagree interval measure details, are shown in Table 1. The 11-item scale, henceforth called the Vertical and Horizontal Internationalization Scale (VHIS), was incorporated into a larger research instrument sent to a sample of senior executives of MNEs with operations in Japan, China, Hong Kong, Singapore, Indonesia and Australia. Accounting for known non-deliverable instruments that were returned by

national post offices, the effective sample was 6,856, although this is likely an overestimate as no undeliverable instruments were returned from Indonesia or China. After eliminating incomplete responses and any firms that prove not to be MNEs, a total of 541 usable responses was received.

= Table 1 here =

Unit non-response bias. To boost overall sample size and to facilitate unit non-response bias checking, a reminder letter and additional questionnaire was sent to 1,200 randomly selected non-responders. This yielded a further useable 99 responses, giving a total of 640, giving a response rate of over 9 percent, comparing well with rates for international business research among senior executives where prior agreement to cooperate has not been sought (Cycyota & Harrison, 2006). In view of the problems associated with response bias estimation models (Vella, 1998), Armstrong and Overton's (1977) procedure for comparing the demographics and mean scores on key variables of early and late responders was followed.

No significant differences between first and second wave respondents were found with respect to broad industry sector (manufacturer, non-manufacturer, $\chi^2 = 2.78$, $p = .10$), country of individual respondent origin (local, non-local to country of firm operation, $\chi^2 = 0.32$, $p = .57$), or language (native, non-native English speaker, $\chi^2 = 1.97$, $p = .16$). Neither were significant differences found with respect either to the summated VI ($t = 1.03$, $p = .30$) or HI ($t = 1.18$, $p = .24$) subscales of the VHIS, or to each subscale's constituent items (VI items: Lower labor prices $t = 1.68$, $p = .09$, Availability of cheaper inputs $t = 1.54$, $p = .12$, Reduction of process/production costs $t = 1.14$, $p = .25$, Procurement of inputs $t = -0.59$, $p = .55$, Lower supply prices $t = 0.20$, $p = .84$; HI items: Expansion of sales $t = 0.67$, $p = .50$, Market opportunities $t = 0.59$, $p = .51$, Market access $t = 1.53$, $p = .13$, Sales to local markets $t = -0.96$, $p = .34$, More selling opportunities $t = 1.75$, $p = .08$), suggesting unit non-response bias not to be a concern.

Item non-response bias. To ensure that the VHIS did not elicit systematic item non-response, all returned questionnaires were examined for potential significant missing-data patterns (Hudson, Seah, Hite, & Haab, 2004). Full responses on VI items were 0.44 percent lower than for HI items. While this percentage is extremely low, it might have resulted from systematic sectoral bias, with some service firms perhaps omitting vertical internationalization items because these items might, at first glance, appear to be of greater relevance to manufacturers. Accordingly, items were tested both individually and summated by VHIS subscale for significant correlations between missing data and firm sector. No significant associations were found for items either individually or summated for firms divided into all service firms versus manufacturers or into just financial services only versus manufacturers, suggesting data were missing completely at random (Hair, Anderson, Tatham, & Black, 1998) and item non-response bias to be absent.

Dimensionality, reliability and discriminant validity. The VHIS was examined using exploratory and confirmatory factor analysis. As Table 2 shows, after excluding the distracter item and using principal component analysis, two discrete and unidimensional components representing the VHIS subscales emerge, together representing 61.09 percent of total variance. The average loading of the VI items on the VI component is .82, whereas their average cross-loading on the HI component is only .12. For the HI subscale items, their average loading on the HI component is .71, while their average cross-loading on the VI component is just .05. The internal consistency reliabilities of both the VI and HI subscales are acceptably above .70, and the correlation between the two subscales' summated mean scores is modest, supporting the subscales' ability to capture two discrete constructs.

= Table 2 here =

Confirmatory factor analyses substantiated the existence of the two discrete and unidimensional constructs of the VHIS, and supported their discriminant validity. Following Hu and

Bentler's (1998) suggestion to employ a combination of fit indices, six were used to obtain a balanced assessment of fit for a two latent variable model accounting both for sample size and number of estimated parameters, and for the sensitivities of different indices to model specification: root-mean-square error of approximation (RMSEA), standardized root-mean-square residual (SRMR), comparative-fit index (CFI), normed fit index (NFI), goodness-of-fit index (GFI), and comparative GFI. Recommended values of RMSEA for acceptable fit are, as they are for all fit indices, subjective and variable, with Browne and Cudeck (1993, p.144) suggesting 'about .08' and MacCallum (1995) highlighting .10 or below as indicating acceptable fit, criteria the VHIS meets (see Table 2). Hu and Bentler (1998, p.449) recommend a cut-off value of .08 for SRMR, a criteria the VHIS meets (Table 2). Values around .90 for CFI, NFI, GFI and adjusted GFI are recommended as indicating acceptable fit, with values close to .95 suggesting good fit (Hu & Bentler, 1998), again criteria the VHIS meets (Table 2). This 2-construct model had a significantly better fit than a single construct confirmatory model ($\Delta\chi^2 = 627.17$, $p < .001$, RMSEA .18, SRMR .10, CFI .74, NFI .73, GFI .77, adjusted GFI .62).

While the above are indicative of good discriminant validity for the VHIS, the direct test of discriminant validity suggested by Bagozzi, Yi and Phillips (1991) was used. This nested model procedure assessed chi square change between when the confirmatory factor model had the parameter between VI and HI unconstrained and then constrained to unity and found strong support for discriminant validity ($\Delta\chi^2 = 299.83$, $p < .001$).

Method variance. To assess the effectiveness in reducing method variance, respectively, of the distracter item, of the main items' construction, and of the overall scale design, both exploratory and confirmatory factor analyses procedures were again used. First, a principal component analysis of the VHIS including the distracter item and specifying a 2-component solution was performed, the

assumption being that the distracter ought, as an item relating to the theoretically discrete construct of oligopolistic reaction (Knickerbocker, 1973), to load only lowly on the VI and HI components. Its loading on VI was .28 and on HI .11, suggesting the distracter item both is unrelated substantively to either VHIS subscale and performs its function well inasmuch as respondents appear to answer both it and VHIS subscales with diligent discrimination. Second, an exploratory principal component analysis was performed with only the VHIS items and the number of components determined by eigenvalues of 1 or higher that produced the 2-component solution shown in Table 2. If the VHIS induces non-diligent responding and therefore high levels of related method variance, two discrete, lowly correlating components would be very unlikely to emerge from exploratory principal component analysis (Harman, 1976).

While these exploratory factor analysis procedures give some assurance that the VHIS induces little method variance, it provides no means of assessing just how much, if any, covariance between the VHIS subscales may in fact exist. Accordingly, a confirmatory factor analysis procedure suggested by Podsakoff, MacKenzie, Lee and Podsakoff (2003, p.894) for partialing out method variance as a latent variable was used. The change in estimated covariance of VI and HI between a confirmatory factor model and the same model but with a latent variable representing method variance was very low, just .002, suggesting negligible effect from method variance (see Table 2).

Study 3. Cross-sample stability

Sample and procedure. To examine the stability of the VHIS in a validation sample, and to provide a larger overall sample for testing its cross-cultural and cross-sectoral stability, it was incorporated into a survey instrument for a different international business study sent to a different sample. This sample comprised 2,894 senior MNE executives throughout the Asia-Pacific region

who had voluntarily assisted with other, unrelated academic research. After a first administration

followed by a reminder to non-responders, 1009 usable responses were obtained, a 35 percent

response rate. Tests for both unit and item non-response bias were insignificant.

Dimensionality, reliability and discriminant validity. Table 2 shows analyses and descriptives

for the VHIS with this new sample. An exploratory principal component analysis excluding the

distracter item again produces two discrete and unidimensional components, this time accounting for

62.55 percent of variance. The average loading of the VI subscale items on the VI component is .83,

and their average cross-loading on the HI subscale component is just .06. For the HI subscale items,

their average item loading on the HI component is .73, while their average cross-loading on the VI

component is only .06. A confirmatory factor analysis again also supports the existence of the two

discrete and unidimensional constructs of the VHIS, with fit indices each proving acceptable. The

internal consistency reliabilities of both the VI and HI subscales are again acceptable, and the

correlation between the two summated mean scores is very modest, once more supporting the

subscales' ability to capture two broadly discrete constructs. A nested structural model test of

discriminant validity found the chi square change between the unconstrained and constrained models

to be significant ($\Delta\chi^2 = 619.47$, $p < .001$).

The *Study 2* and *Study 3* samples were pooled to produce a full sample of 1,649. Analyses of

this pooled sample also reinforce the dimensionality and reliability of the VHIS (see Table 2), with a

structural model test of discriminant validity finding the chi square change between the

unconstrained and constrained models significant ($\Delta\chi^2 = 507.13$, $p < .001$).

Cross-national stability. The equivalence of measures across countries and cultures is of

particular importance in international business research (Harzing et al., 2009; Ralston, Russell, &

Egri, 2018). Cross-group confirmatory factor analysis was used to test some degree of cross-cultural

factorial equivalence of the VHIS by splitting the full sample into native English speakers and non-native English speakers based on respondents' personal as opposed to their MNEs' national origin.

When the factorial stability between native and non-native English speakers is tested, the change in model chi square across the two groups between the unconstrained and constrained models' measurement weights is insignificant ($\Delta\chi^2 = 8.35$, $p = .40$), indicating factorial invariance. While the division of the sample into native and non-native English speakers is a broad aggregation, it suggests that the scale retains its measurement properties across nationalities at a general level of aggregation, hinting both at the appropriateness of its use with international samples in English and its possible suitability for translation into other languages.

Cross-sectoral stability. For the VHIS to be of practical use across the gamut of possible international business research applications, it must exhibit validity and stability across industry sectors. The metrical properties of the VHIS across manufacturing and service firms were therefore examined. The dimensionality and reliabilities of the VI and HI subscales remain adequate for each broad industry sector (see Table 3). When factorial invariance between manufacturers and service firms is tested, the change in model chi square between unconstrained and constrained models' measurement weights is insignificant ($\Delta\chi^2 = 12.98$, $p = .11$), supporting the cross-sectoral stability of the VHIS at a broad level of aggregation.

<p style="text-align:center">= Table 3 here =</p>

Study 4. Temporal stability

To examine the temporal stability of the VHIS, a retest instrument was administered to 200 firms selected randomly from the *Study 2* respondents. The retest was administered 12 month after the initial validation and resulted in 115 verifiably matched responses. For this sample, the internal

consistency reliabilities are broadly similar (VI subscale initial test .90, retest, .83; HI subscale initial test .77, retest .78). The mean scores for each of the subscales are not significantly different (VI subscale test mean 4.21, SD 1.01, retest mean 4.15, SD 0.99, $t = 1.31$, $p = .19$; HI subscale test mean 5.08, SD 0.62, retest mean 5.13, SD 0.61, $t = -0.90$, $p = .37$). The test-retest coefficient of reliability for the VI subscale is .73 ($p < .0001$), for the HI subscale, .55 ($p < .0001$), suggesting 'exemplary' long-run temporal stability by the criteria suggested by Robinson, Shaver and Wrightsman (1991, p.13). The factorial stability between initial test and retest responses was examined using the cross-group structural model invariance procedure suggested by Byrne (2001). The change in model chi square across the test and retest responses between the unconstrained and constrained models' measurement weights is insignificant ($\Delta\chi^2 = 5.83$, $p = .44$). Taken together, the similar internal reliabilities, insignificant subscale mean differences, the good test-retest correlations, and the factorial invariance of the VHIS over a 12-month period suggest excellent temporal stability.

Study 5. Country-specific question stem

All the validation analyses reported above are based on a question stem about general internationalization rather than about internationalization in a specific country. To examine the dimensionality of the VHIS using a country-specific question stem, it was incorporated into an instrument for a study of MNEs in relation to a particular country, China, with this wording: *Thinking generally, in deciding to have direct investments in China, do you agree that the following are important motivating factors for your firm?* The obtained sample of 1,125 MNEs comes from a total of 36 different countries, each MNE having either or both processing/production or sales/marketing direct investments in China.

Table 4 shows results of exploratory and confirmatory factor analyses for the whole sample and by broad industry sector, with both types of analyses supporting the dimensionality of the VHIS in each sector. Also shown are correlations between the VI and HI subscales. These are low in each case, and insignificant except in the instance of manufacturing MNEs, again lending support to the dimensionality of the VHIS and its ability to discriminate between VI and HI. A direct, nested structural model test of discriminant validity found the chi square change between the unconstrained and constrained models to be significant ($\Delta\chi^2 = 656.42$, $p < .001$). The reliabilities of the subscales are adequate for each sample.

A test of cross-sectoral factorial invariance was significant ($\Delta\chi^2 = 21.02$, $p = .007$). To investigate why this should be so, following Byrne (2001) the parameter constraint on each VI and HI item was relaxed in turn with the result that the HI item Market Access was found to have a measurement weight that was not invariant across broad sectors. The contribution of this item to the Cronbach alpha of the HI subscale was examined on a sector by sector basis and found to make a positive contribution to each sector except to the retail, restaurants and hotels sector. When this sector is excluded from a factorial invariance test across broad manufacturing and service groups, the change in chi square drops substantially and is insignificant at a 1 percent level ($\Delta\chi^2 = 15.67$, $p = .043$). The factorial invariance across native and non-native English speakers was also tested and no significant difference between the constrained and unconstrained models was found ($\Delta\chi^2 = 6.29$, $p = .62$), indicating that the VHIS remains cross-culturally stable with a country-specific question stem.

= Table 4 here =

Criterion-related validity. Establishing criterion-related validity is notoriously problematic because, in Nunnally and Bernstein's words, 'obtaining a good criterion may actually be more

difficult than obtaining a good predictor' (1994, p.96). Certainly this is the case for VI and HI with respect to convergent validity due to a lack of any pre-existing specific measures for VI and HI and, in terms of potentially plausible proxy measures, the non-availability of adequate firm-level data on proportion of sales to host countries by MNE subsidiaries. However, the predictive validity of the VHIS for both the general and country-specific question-stem samples was examined with respect to variables that the VI and HI subscales might be predicted to correlate with.

Objective variables. Data were collected in relation to three objective variables, industry sector, extent of internationalization, and size.

Industry sector. VI and HI might reasonably be predicted to correlate differently with firms by industry sector. For example, manufacturers' FDI might be anticipated to correlate positively with VI but perhaps negatively with HI, especially in the case of manufacturers of industrial products where competition on price of relatively commoditized products, and therefore constant efforts to increase supply efficiencies, are more common than in some other sectors. Conversely, service firms' FDI, perhaps most particularly business financial service firms that have relatively small back-office functions compared to other service sectors, might be anticipated to correlate positively with HI but negatively with VI. Such relationships might be expected to hold both on a general and a country-specific question-stem basis.

Accordingly, dummy variables were created for industrial product manufacturers and business financial services firms respectively, and significant correlations in the directions anticipated were found for both general and country-specific samples, lending support to the criterion-related validity of the VHIS (see Table 5).

Internationalization extent. The number of countries in which an MNE operates might be predicted to correlate differently with VI and HI. On one hand, because HI reflects a strategy of exploiting firm-specific assets in foreign markets the logical conclusion of which is sales in every

viable market world-wide, the number of countries in which an MNE has substantial marketing and sales operations might reasonably be hypothesized to correlate positively with HI, but either not to correlate significantly or perhaps even to correlate negatively with VI. On the other hand, because VI represents the performance of only so many processing and production activities as are necessary in only the limited number of countries where it makes economic sense, the relationship between the number of countries where an MNE has processing and production operations for both and VI and HI is less obvious, with only very small or no significant correlations predictable. Using simple counts of countries where MNEs have, respectively, marketing/sales operations and processing/production operations to measure internationalization extent (Lu & Beamish, 2004), the patterns of correlations suggested above were found for both general and country-specific samples (see Table 5), again suggestive of criterion-related validity.

MNE size. Again, because HI represents primarily a strategy of firm-specific asset exploitation in foreign markets, MNE size might be predicted to correlate positively with HI. However, because VI represents primarily a strategy of increasing supply efficiencies, there is no compelling logic to suggest any correlation between MNE size and VI. Using number of employees as a proxy for size, these anticipated correlations were borne out in both general and country-specific samples (see table 5), suggestive again of criterion-related validity.

Subjective variables. Because HI is motivated by a desire to exploit firm-specific assets, HI ought positively to correlate with advantageous foreign market conditions, whereas VI ought not to. Alternatively, because VI is motivated by the desire to exploit advantageous supply circumstances, VI ought positively to correlate with measures of low-cost foreign supply circumstances, while HI ought not to.

To measure advantageous market conditions, a 4-item measure asking how well or how badly China scored in terms of attractiveness as an FDI location with respect to (i) gross domestic

product growth, (ii) absolute market size, (iii) market size per capita, and (iv) market size increase

per capita was created and then administered to the country-specific sample. The same items were

also administered to the general, non-country-specific sample, but with a general rather than China-

specific question stem. To measure low-cost supply circumstances, a 4-item measure of how well or

how badly China scored in terms of attractiveness as an FDI location with respect to costs (i) of

employees, (ii) of commercial space, (iii) of business supplies, and (iv) of general operating

expenses was created and adapted accordingly for administration to the country-specific and general

samples. The items of each of these 4-item measures were assessed using a 6-point interval measure

running from *very badly* to *very well*.

For the country-specific question-stem respondents, each predictive measure was

incorporated into an instrument sent to all 1,125 country-specific respondents (Study 5) a full eight

months after they had completed the VHIS, so effects of potential method variance were

substantially ameliorated by temporal distance (Podsakoff et al., 2012). Some 483 useable and

matchable responses were received. For the general question stem respondents, an instrument

containing the two predictive validity measures was sent to a randomly selected sample of 400 who

had completed the VHIS in its general rather than country-specific format. This instrument was sent

two months after respondents had completed the VHIS, thereby diminishing somewhat the effects of

method variance (Podsakoff et al., 2012). Some 221 fully completed and matchable responses were

received. The advantageous market conditions measure's Cronbach's alpha for the general sample

was .74, for the country-specific sample .81, and respective statistics for the low-cost supply

circumstances measure were .88 and .81. Correlations between the VHIS's VI and HI subscales with

the respective criterion variables are shown in Table 5 and are in accordance with expectations,

thereby lending support to the criterion-related validity of the VHIS.

= Table 5 here =

Classificatory ability

The classificatory ability of the VHIS was examined with the country-specific question-stem

sample using non-hierarchical cluster analysis, specifying a non-running mean procedure to avoid

sensitivity to peculiarities arising from initial case order effects (Hair et al., 1998). Using the

summated mean scores of the VI and HI subscales as the clustering variables, logically a 3-cluster

solution should produce a meaningful and discriminating set of clusters, with one cluster being

MNEs with FDI in China representing predominantly VI, another cluster being MNEs with China

FDI representing predominantly HI, and a further cluster being MNEs with China FDI representing a

combination of both VI and HI. As the sample comprises MNEs specifically with either or both

processing/production or sales/marketing FDI in China, a fourth theoretically possible cluster

comprising MNEs with China FDI constituting neither VI or HI cannot logically be anticipated.

In analysis results found the predominantly VI cluster comprises 174, 15.47 percent, of the

MNEs (VI mean score 4.59, HI mean score 3.82), the predominantly HI cluster comprises 297,

26.40 percent, of MNEs (VI mean score 2.80, HI mean score 5.14), and the combined VI and HI

cluster comprises 654 MNEs, 58.13 percent (VI mean score 4.60, HI mean score 5.29). As the three

anticipated clusters emerge, good classificatory ability for the VHIS is suggested.

Discussion

The need for valid and reliable measures of key constructs used in theorizing and empirically testing

propositions about MNE strategy, structure and performance is highlighted by several scholars

(Marano, Arregle, Hitt, Spadafora, & van Essen, 2016; Oesterle & Richta, 2013; Verbeke &

Brugman, 2009; Wiersema & Bowen, 2011). Our paper seeks to draw attention to the need to

develop and validate firm-level measures of the VI and HI constructs to help facilitate theorization

and testing of propositions about MNEs and the extent to which their internationalization constitutes

VI and HI. Our series of studies using senior MNE manager informants across a range of sectors and countries, has enabled us to develop and validate a research-practical and metrically robust measurement tool to render hypotheses concerning VI and HI empirically testable for both MNEs' overall and for their country-specific FDI. The resultant VHIS's VI and HI subscales prove to be content-valid, largely independent of each other, unidimensional, possessed of discriminant validity, internally reliable, temporally stable, generalizable across populations, cross-culturally and cross-sectorially valid, possessed of good criterion-related validity using both objective and subjective measures, and capable of categorizing MNEs by the varying extents to which their FDI constitutes VI and HI.

Limitations and further research

Although this study uses relatively large samples and directly addresses cross-population validity, further validation with different samples would be useful. In terms of the cross-national validity of the VHIS, the examination possible in this research uses only groupings of native and non-native English speakers. While this is appropriate to a highly diverse international sample and is in keeping with the realities of the nature of many senior MNE manager field-samples, further cross-national validity testing on a national-origin basis would be useful. Additionally, for some samples, the VHIS will likely need translation into languages other than English, and certainly further research in this regard can usefully examine whether or not the VHIS's metrical properties are robust to translation. Further, this research seeks only to examine the VHIS in relation to general and country-specific FDI, but does not investigate the VHIS in relation to individual operations within countries. Further research to assess the VHIS in this regard is needed as FDI for some firms in some countries comprises more than one operation, some of which are possibly VI, while others are possibly HI.

Notwithstanding the need for additional research to further validate the VHIS that attends the

development of any new metric, the VHIS now offers a means by which consideration of VI and HI

can be operationalized in a range of research questions with direct practical relevance to managers in

both business and public policy sectors.

Performance and organization. In terms of performance, it might reasonably be anticipated

that the degree to which an MNE's general and country-specific FDI constitutes either VI or HI will

have some effect because the scope for profit enhancement from cost reduction compared to market

expansion is heterogeneous. With respect to the effect of internationalization extent on performance,

controlling for the degree to which such internationalization constitutes VI or HI will allow a more

finely nuanced analysis in this already heavily but inconclusively researched area (Cardinal, Miller,

& Palich, 2011; Kirca, Roth, Hult, & Cavusgil, 2012; Li, 2007; Matysiak & Bausch, 2012; Nguyen,

2017). Similarly, incorporating consideration of the extent to which an MNE's FDI is VI or HI will

facilitate finer-grained analysis of entry mode and governance structures. For instance, the merits of

acquisitions versus green-field entry might be predicted to differ for VI and HI, as can governance

and organization structures. For example, FDI comprising VI is likely to demand relatively close

parental control in view of the need for efficient integration into an MNE's production chain,

whereas FDI comprising HI might be predicted to reflect governance and organizational structures

that facilitate mandates for greater autonomy in order that responsiveness to local markets can be

achieved. Research to derive nuanced answers to these performance and organization issues can help

inform business managers better to align their degrees of VI and HI towards performance

optimization.

Host-country effects. An extensive literature exists on the economic growth and wider

political economy effects of MNEs' investments on host countries (Abdouli & Hammami, 2017;

Iamsiraroj, 2016; Gunby, Jin, & Reed, 2017; Poon & Thompson, 1998; Temiz & Gökmen, 2014),

ranging from technology transfers to local firms (Fan, Cui, Li, & Zhu, 2016; Nuruzzaman, Singh, &

Pattnaik, in press; Saranga, Schotter, & Mudambi, in press; Thompson, 2002, 2003; Wu, Ma, &

Zhuo, 2017), to the pressures, both formal and informal, that relatively footloose MNEs exert on

hosts' political and institutional systems as their governments seek first to attract and then retain FDI

(Banerjee, Venaik, & Brewer, in press; Büthe & Mattli, 2011; Jensen, 2003; Perkins, 2010;

Rizopoulos & Sergakis, 2010; Sun, Mellahi, & Thun, 2010; Thompson & Poon, 2000). As

Beugelsdijk et al. (2008) point out, much of this literature currently fails to distinguish between FDI

arising from VI and HI. While host-country in aggregate data for US MNEs exists that allow some

limited theory testing, more broadly nuanced source-country, industry and firm-level analyses have

hitherto been constrained because no appropriate firm-level measures of VI and HI have existed. In

consequence, not only are the likely heterogeneous externalities of VI and HI as yet unknown in

detail, but the benefits, and costs, of FDI to host countries are analyzed only at an aggregate level.

As a result public policy managers in host governments cannot develop appropriately nuanced policy

to attract, or indeed repel, VI or HI according to wider political economy goals. The same also

applies to both national government and international organization managers that seek to assist

economic development in emerging economies.

Location choices. Although Dunning's (1998) lament that location choice is a central but

neglected factor in MNE research has been partially addressed in subsequent research, the differing

location determinants of FDI constituting, respectively, VI and HI still remain somewhat

overlooked, attaining only brief mention in recent literature surveys (Kim & Aguilera, 2016; Li,

Quan, Stoian, & Azar, 2018; Nielsen, Asmussen, & Weatherall, 2017). Naturally, other factors, such

as industry and competitive and corporate strategy, will condition the effect on MNEs' location

choices for their VI and HI, but it might be hypothesized, for example, that VI will require

institutional and political stability in a location more than HI due to the more critical nature of VI to

the ability of an MNE to produce its goods or services for all its markets. Whereas VI may represent sunk costs that need to operate effectively and continuously to ensure the successful operation of an MNE's entire value chain, HI in any particular location is more likely to be one of several foreign market expansions. Consequently, HI to any single country represents a less critical firm-wide risk, meaning that institutional and political stability become less critical requirements for HI than for VI. The nature of pre-existing supplier concentrations and infrastructural quality in FDI host-countries will also likely be different for VI and HI. For instance, VI might be anticipated to require in host locations numerous and reliable supplies of materials and components, whereas HI might be expected to require infrastructure and services associated with marketing and sales activities. In helping to make examination of FDI location determinants more empirically tractable, the development of the VHIS can assist not just business managers to make more informed FDI location decisions, but can also help public policy managers to develop policies to attract, or deter, VI and HI in line with broader political economy objectives.

References

Abdouli, M. & Hammami, S. (2017). Investigating the causality links between environmental quality, foreign direct investment and economic growth in MENA countries. *International Business Review*, *26*, 264-278.

Aizenman, J. & Marion, N. (2004). The merits of horizontal versus vertical FDI in the presence of uncertainty. *Journal of International Economics*, *62*, 125-148.

Alfaro, L. & Charlton, A. (2009). Intra-industry foreign direct investment. *American Economic Review*, *99*, 2096-2119.

Armstrong, M. C. &. Overton, T. S. (1977). Estimating nonresponse bias in mail surveys. *Journal of Marketing Research*, *14*, 396-402.

Aronsson, T. & Koskela, E. (2011). Redistributive income taxation, outsourcing and foreign direct investment. *Economic Letters*, *110*, 86-89.

Azémar, C. & Desbordes, R. (2010). Short-run strategies for attracting foreign direct investment. *The World Economy*, *33*, 928-957.

Bagozzi, R. P., Yi, Y., & Phillips, L. W. (1991). Assessing Construct Validity in Organizational Research. *Administrative Science Quarterly*, *36*, 421-458.

Baldwin, R. & Okubo, T. (2014). Networked FDI: sales and sourcing patterns of Japanese foreign affiliates. *The World Economy*, *37*, 1051-1080.

Banerjee, S., Venaik, S., & Brewer, P. (in press). Analysing corporate political activity in MNC subsidiaries through the integration-responsiveness framework. *International Business Review*.

Barbopoulos, L., Marshall, A., MacInnes, C., & McColgan, P. (2014). Foreign direct investment in emerging markets and acquirers' value gains. *International Business Review*, *23*, 604-619.

BEA (2017). *Direct investment and multinational enterprises: comprehensive data*. Available online at:

https://www.bea.gov/international/direct_investment_multinational_companies_comprehensive_data.htm

Benito, G. (2015). Why and how motives (still) matter. *Multinational Business Review*, *23*, 15-24.

Beugelsdijk, S., Smeets, R., & Zwinkels, R. (2008). The impact of horizontal and vertical FDI on host's country economic growth. *International Business Review*, *17*, 452-472.

Bollen, K. A. & Diamantopoulos, A. (2017). In defense of causal-formative indicators: a minority report. *Psychological Methods*, *22*, 581-596.

Braconier, H., Norbäck, P. J., & Urban, D. (2005). Multinational enterprises and wage costs: vertical FDI revisited. *Journal of International Economics*, *67*, 446-470.

Brown, C. L., Cavusgil, T., & Lord, A. W. (2015). Country-risk measurement and analysis: a new conceptualization and managerial tool. *International Business Review*, *24*, 246-265.

Browne, M. W. & Cudeck, R. (1993). Alternative ways of assessing model fit. In Bollen, K. A. & Long, J. S. (eds.) *Testing Structural Equation Models*. Newbury Park: Sage Publications, 136-162.

Buckley, P. J., Clegg, J., Cross, A. R., Liu, X., Voss, H., & Zheng, P. (2007). The determinants of Chinese outward foreign direct investment. *Journal of International Business Studies*, *38*, 499-518.

Büthe, T. & Mattli, W. (2011). *The new global rulers: the privatization of regulation in the world economy*. Princeton, NJ: Princeton University Press.

Byrne, B. M. (2001). *Structural equation modeling with AMOS*. Mahwah, NJ: Lawrence Erlbaum.

Cardinal, L. B., Miller, C. C., & Palich, L. E. (2011). Breaking the cycle of iteration: forensic failures of international diversification and firm performance research. *Global Strategy Journal*, *1*, 175-186.

Caves, R. E. (1971). International corporations: the industrial economics of foreign investment. *Economica*, *38*, 1-27.

Caves, R. E. (1996). *Multinational Enterprise and Economic Analysis*, 2nd Edition. Cambridge: Cambridge University Press.

Caves, R. E. (2007). *Multinational Enterprise and Economic Analysis*, 3rd Edition. Cambridge: Cambridge University Press.

Chen, H. & Chen, T. (1998). Network linkages and location choice in foreign direct investment. *Journal of International Business Studies*, *29*, 445-467.

Christoffersen, J., Plenborg, T., & Robson, M. J. (2014). Measures of strategic alliance performance, classified and assessed. *International Business Review*, *23*, 479-489.

Cieślik, A. (2017). Determinants of MNE activity in Poland: the case of firms from EU-15. *Entrepreneurial Business and Economics Review*, *5*, 151-167.

Coltman, T., Divinney, T. M., Midgley, D. F., & Venaik, S. (2008). Formative versus reflective measurement models: two applications of formative measurement. *Journal of Business Research*, *61*, 1250-1262.

Cycyota, C. S. & Harrison, D. A. (2006). What (not) to expect when surveying executives: a meta-analysis of top manager response rates and techniques over time. *Organizational Research Methods*, *9*, 133-160.

Davies, R. B. (2008). Hunting high and low for vertical FDI. *Review of International Economics*, *16*, 250-267.

Di Gregorio, D., Musteen, M., & Thomas, D. E. (2008). International new ventures: the cross-border nexus of individuals and opportunities. *Journal of World Business, 43*, 186-196.

Dow, D. & Ferencikova, S. (2010). More than just national cultural distance: testing new distance scales on FDI in Slovakia. *International Business Review*, *19*, 46-58.

Dunning, J. H. (1998). Location and the multinational enterprise: a neglected factor? *Journal of International Business Studies*, *29*, 45-66.

Fan, D., Cui, L., Li, Y., & Zhu, C. J. (2016). Localized learning by emerging multinational enterprises in developed host countries: a fuzzy-set analysis of Chinese foreign direct investment in Australia. *International Business Review, 25*, 187-203.

Feenstra, R. C. & Hanson, G. H. (1997). Foreign direct investment and relative wages: evidence from Mexico's maquiladoras. *Journal of International Economics*, *42*, 371-393.

Finn, A. & Wang, L. (2014). Formative vs. reflective measures: facets of variation. *Journal of Business Research*, *67*, 2821-2826.

Fukao, K. & Wei, Y. (2008). How do the location determinants of vertical and horizontal FDI differ? Hitotsubashi University Research Unit for Statistical Analysis in Social Sciences, Discussion paper 233, 1-45.

Galan, J. I., Gonzalez-Benito, J., & Zuñiga-Vicente, J. A. (2007). Factors determining the location decisions of Spanish MNEs: an analysis based on the investment development path. *Journal of International Business Studies*, *38*, 975-997.

Gerschewski, S. & Xiao, S. S. (2015). Beyond financial indicators: an assessment of the measurement of performance for international new ventures. *International Business Review*, *24*, 615-629.

Gunby, P., Jin, Y., & Reed, R. (2017). Did FDI really cause Chinese economic growth? A meta-analysis. *World Development*, *90*, 242-255.

Hair, J. F., Anderson, R. L., Tatham, R. L., & Black, W. C. (1998). *Multivariate Data Analysis*. Upper Saddle River: Prentice Hall.

Hardin, A. (2017). A call for theory to support the use of causal-formative indicators: a commentary on Bollen and Diamantopoulos (2017). *Psychological Methods*, *22*, 597-604.

Harman, H. H. (1976). *Modern Factor Analysis*, 3rd Edition. Chicago, IL: University of Chicago Press.

Harzing, A., Baldueza, J., Barner-Rasmussen, W., Barzantny, C., Canabal, A., Davila, A., Espejo, A., Ferreira, R., Giroud, A., Koester, K., Liang, Y., Mockaitis, A., Morley, M. J., Myloni, B., Odusanya, J. O. T., O'Sullivan, S. L., Palaniappan, A. K., Prochno, P., Choudhury, S. R., Saka-Helmhout, S., Siengthai, S., Viswat, L., Soydas, A. U., & Zander, L. (2009). Rating versus ranking: what is the best way to reduce response and language bias in cross-national research? *International Business Review, 18*, 417-432.

Helpman, E. M. (1984). A simple theory of trade with multinational corporations. *Journal of Political Economy*, *92*, 451-471.

Hu, L. T. & Bentler, P. M. (1998). Fit indices in covariance structure modeling: sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*, 424-453.

Huett, P., Baum, M., Schwens, C., & Kabst, R. (2014). Foreign direct investment location choice of small- and medium-sized enterprises: the risk of value erosion of firm-specific resources. *International Business Review*, *23*, 952-965.

Hudson, D., Seah. L. H., Hite, D., & Haab, T. (2004). Telephone presurveys, self-selection, and non-response bias to mail and internet surveys in economic research. *Applied Economics Letters*, *11*, 237-240.

Iamsiraroj, S. (2016). The foreign direct investment–economic growth nexus. *International Review of Economics & Finance*, *42*, 116-133.

Jabbour, L. (2010). Offshoring and firm performance: evidence from French manufacturing industry. *World Economy*, *33*, 507-524.

Jensen, N. M. (2003). Democratic governance and multinational corporations: political regimes and inflows of foreign direct investment. *International Organization*, *57*, 587-616.

Jiang, G., Holburn, G. L. F., & Beamish, P. W. (in press). Repeat market entries in the internationalization process: the impact of investment motives and corporate capabilities. *Global Strategy Journal*.

Kim, J. U. & Aguilera, R. V. (2016). Foreign location choice: review and extensions. *International Journal of Management Reviews*, *18*, 133-159.

Kinda, T. (2013). Beyond natural resources: horizontal and vertical FDI diversification in Sub-Saharan Africa. *Applied Economics*, *45*, 3587-3598.

Kirca, A. H., Roth, K., Hult, T. M., & Cavusgil, S. T. (2012). The role of context in the multinationality-performance relationship: a meta-analytic review. *Global Strategy Journal*, *2*, 108-121.

Knickerbocker, F. T. (1973). *Oligopolistic Reaction and Multinational Enterprise*. Boston: Division

    of Research, Graduate School of Business Administration, Harvard University.

Kukulski, J. & Ryan, M. (2011). Investment history and market orientation effects in the TFP-FDI

    relationship. *The World Economy*, *34*, 546-567.

Lankhuizen, M. (2014). The (im)possibility of distinguishing horizontal and vertical motivations for

    FDI. *Review of Development Economics*, *18*, 139-151.

Li, L. (2007). Multinationality and performance: a synthetic review and research agenda.

    International Journal of Management Reviews, *9*, 117–139.

Li, X., Quan, R., Stoian, M., & Azar, G. (2018). Do MNEs from developed and emerging economies

    differ in their location choice of FDI? A 36-year review. *International Business Review*, *27*,

    1089-1103.

Li, Q., Scollay, R., & Gilbert, J. (2017). Analyzing the effects of the regional comprehensive

    economic partnership on FDI in a CGE framework with firm heterogeneity. *Economic

    Modelling*, *67*, 409-420.

Lu, J. W. & Beamish, P. W. (2004). International diversification and firm performance: the S-curve

    hypothesis. *Academy of Management Journal*, *47*, 598-609.

MacCallum, R. C. (1995). Model specification: procedures, strategies, and related issues. In Hoyle

    R. H. (Ed.) *Structural equation modeling: concepts, issues and applications*. Newbury Park,

    CA: Sage, 76-99.

Marano, V., Arregle, J., Hitt, M. A., Spadafora, E., & van Essen, M. (2016). Home country

    institutions and the internationalization-performance relationship: a meta-analytic review.

    Journal of Management, *42*, 1075-1110.

Markusen, J. R. (1984). Multinational firms, multi-plant economies, and gains from trade. *Journal of

    International Economics*, *16*, 205-226.

Markusen, J. R. & Maskus, K. E. (2002). Discriminating among alternative theories of the

    multinational enterprise. *Review of International Economics*, *10*, 694-707.

Markusen, J. R. & Venables, A. J. (1998). Multinational firms and the new trade theory. *Journal of International Economics*, *46*, 183-203.

Markusen, J. R. & Venables, A. J. (2000). The theory of endowment, intra-industry and multi-national trade. *Journal of International Economics*, *52*, 209-234.

Matysiak, L. & Bausch, A. (2012). Antecedents of MNE performance: blinded by the obvious in 35 years of literature. *Multinational Business Review*, *20*, 178-211.

Mesquita, L. F. (2016). Location and the global advantage of firms. *Global Strategy Journal*, *6*, 3-12.

METI (2016). *Basic Survey on Overseas Business Activities*. Tokyo: Ministry of Economy, Trade and Industry.

Nachum, L. & Zaheer, S. (2005). The persistence of distance? The impact of technology on MNE motivations for foreign investment. *Strategic Management Journal*, *26*, 747-767.

Nam, D., Kim, J., Arthurs, J. D., Sosik, J. J., & Cullen, J. B. (2016). Measurement and structural invariance of entrepreneurial investment climate: a cross-country scale development. *International Business Review*, *25*, 1053-1065.

Nes, E. B., Yelkur, R., & Silkoset, R. (2014). Consumer affinity for foreign countries: construct development, buying behavior consequences and animosity contrasts. *International Business Review*, *23*, 774-784.

Nguyen, Q. T. K. (2017). Multinationality and performance literature: a critical review and future research agenda. *Management International Review*, *57*, 311-347.

Nielsen, B. B., Asmussen, C. G., & Weatherall, C. D. (2017). The location choice of foreign direct investments: empirical evidence and methodological challenges. *Journal of World Business*, *52*, 62-82.

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

Nuruzzaman, N., Singh, D., & Pattnaik, C. (in press). Competing to be innovative: foreign

competition and imitative innovation of emerging economy firms. *International Business

Review*.

Oesterle, M. & Richta, H. N. (2013). Internationalisation and firm performance: state of empirical

research efforts and need for improved approaches. *European Journal of International

Management*, *7*, 204-224.

Pappu, R. & Quester, P. (2010). Country equity: conceptualization and empirical evidence.

International Business Review, *19*, 276-291.

Perkins, J. (2010). *The secret history of the American empire.* New York: Plume Publishing.

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases

in behavioral research: a critical review of the literature and recommended remedies. *Journal

of Applied Psychology*, *88*, 879-903.

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social

science research and recommendations on how to control it. *Annual Review of Psychology,

63*, 539-569.

Poon, J. P. H. & Thompson, E. R. (1998). Foreign direct investment and economic growth: evidence

from Asia and Latin America. *Journal of Economic Development*, *23*, 141-160.

Ralston, D. A., Russell, C. J., & Egri, C. P. (2018). Business values dimensions: a cross-culturally

developed measure of workforce values. *International Business Review*, *27*, 1189-1199.*.*

Rizopoulos, A. & Sergakis, D. E. (2010). MNEs and policy networks: institutional embeddedness

and strategic choice. *Journal of World Business*, *45*, 250-256.

Robinson, J. P., Shaver, P. R., & Wrightsman, L. S. (1991). Criteria for scale selection and

evaluation. In Robinson, J. P., Shaver, P. R., & Wrightsman, L. S. (Eds.), *Measures of

Personality and Social Psychological Attitudes*. San Diego: Academic Press.

Saranga, H., Schotter, P. J. S., & Mudambi, R. (in press). The double helix effect: catch-up and

local-foreign co-evolution in the Indian and Chinese automotive industries. *International

Business Review*.

Schriesheim, C. A., Powers, K. J., Scandura, T. A., Gardiner, C. C., & Lankau, M. J. (1993). Improving construct measurement in management research. *Journal of Management*, *19*, 385-417.

Singla, C. & George, R. (2013). Internationalization and performance: a contextual analysis of Indian firms. *Journal of Business Research*, *66*, 2500-2506.

Slangen, A. H. L. & Beugelsdijk, S. (2010). The impact of institutional hazards on foreign multinational activity: a contingency perspective. *Journal of International Business Studies*, *41*, 980-995.

Slangen, A. H. L. & van Tudler, R. J. M. (2009). Cultural distance, political risk, or governance quality? Towards a more accurate conceptualization and measurement of external uncertainty in foreign entry mode research. *International Business Review, 18*, 276-291.

Stack, M. M., Ravishankar, G., & Pentecost, E. (2017). Foreign direct investment in the eastern European countries: determinants and performance. *Structural Change and Economic Dynamics*, *41*, 86-97.

Sullivan, D. (1994). Measuring the degree of internationalization of a firm. *Journal of International Business Studies*, *25*, 325-342.

Sun, P., Mellahi, K., & Thun, E. (2010). The dynamic value of MNE political embeddedness: the case of the Chinese automobile industry. *Journal of International Business Studies*, *41*, 1161-1182.

Swedish Research Institute of Industrial Economics (2003). *Foreign operations of Swedish manufacturing firms*. Available online at: https://www.ifn.se/eng/research/databases_10/foreign_operations_of_swedish_manufacturing_firms

Tanaka, K. (2017). Foreign direct investment and temporary workers in Japan. *Journal of Asian Economics*, *48*, 87-99.

Temiz, D. & Gökmen, A. (2014). FDI inflow as an international business operation by MNCs and economic growth: an empirical study on Turkey. *International Business Review*, *23*, 145-154.

Thompson, E. R. (2002). Clustering of foreign direct investment and enhanced technology transfer: evidence from Hong Kong firms in China. *World Development*, *30*, 873-889.

Thompson, E. R. (2003). Technology transfer to China by Hong Kong's cross-border garment firms. *The Developing Economies*, *41*, 88-111.

Thompson, E. R. & Phua, F. T. T. (2005). Reliability among senior managers of the Marlowe-Crowne short-form social desirability scale. *Journal of Business and Psychology*, *19*, 541-554.

Thompson, E. R. & Poon, J. P. H. (2000). ASEAN after the financial crisis: links between foreign direct investment and regulatory change, *ASEAN Economic Bulletin*, *17*, 1-14.

Vella, F. (1998). Estimating models with sample selection bias: a survey. *The Journal of Human Resources*, *33*, 127-168.

Venables, A. J. (1999). Fragmentation and multinational production. *European Economic Review*, *43*, 935-945.

Verbeke, A. & Brugman, P. (2009). Triple testing the quality of multinationality-performance research: an internalization theory perspective. *International Business Review*, *18*, 265–275.

Wacker, K. M. (2016). (When) Should we use foreign direct investment data to measure the activities of multinational corporations? Theory and evidence. *Review of International Economics*, *24*, 980-999.

Waldkirch, A. (2011). Comparative advantage FDI? A host country perspective. *Review of World Economics*, *147*, 485-505.

Wiersema, M. F. & Bowen, H. P. (2011). The relationship between international diversification and firm performance: why it remains a puzzle. *Global Strategy Journal*, *1*, 152-170.

Wu, J., Ma, Z., & Zhuo, S. (2017). Enhancing national innovative capacity: the impact of high-tech

international trade and inward foreign direct investment. *International Business Review*, *26*,

502-514.

Yeaple, S. R. (2003). The complex integration strategies of multinationals and cross country

dependencies in the structure of foreign direct investment. *Journal of International*

Economics, *60*, 293-314.

Table 1. Vertical and Horizontal Internationalization Scale (VHIS) [a,b,c]

1. Expansion of sales

2. Lower labor prices

3. Market opportunities

4. Activities of close competitors[d]

5. Availability of cheaper inputs

6. Market access

7. Reduction of process/production costs

8. Sales to local markets

9. Procurement of inputs

10. More selling opportunities

11. Lower supply prices

(a). Questions: (i) General question stem, Thinking generally, in deciding to have direct investments in one or more foreign economy, do you agree that the following are important motivating factors for your firm? (ii) Country-specific question stem, Thinking generally, in deciding to have direct investments in [country name], do you agree that the following are important motivating factors for your firm?

(b). To encourage full answering of all items from firms of all sectors, this exhortation was added after the question stem: Whether a service or manufacturing firm, please complete ALL items in this question.

(c). Interval measure: 1 = Strongly Disagree, 2 = Disagree, 3 = Slightly Disagree, 4 = Slightly Agree, 5 = Agree, 6 = Strongly Agree.

(d). Distracter item to reduce method variance and which is not included in scale calculations.

Table 2. VHIS factor analyses, reliabilities and descriptives

	Study 2 Sample N = 640		Study 3 Sample N = 1009		Pooled Study 2 and 3 Sample. N = 1649	
	VI subscale[a]	HI subscale[a]	VI subscale	HI subscale	VI subscale	HI subscale
Principal component item loadings[b, c]						
7. Reduction of process/production costs	**.87**	.15	**.86**	.03	**.87**	.08
11. Lower supply prices	**.87**	.14	**.89**	.03	**.89**	.07
5. Availability of cheaper inputs	**.85**	.11	**.83**	.07	**.84**	.09
2. Lower labor prices	**.78**	.03	**.80**	-.02	**.79**	-.01
9. Procurement of inputs	**.75**	.18	**.78**	.17	**.77**	.17
10. More selling opportunities	.08	**.78**	.12	**.80**	.11	**.79**
3. Market opportunities	.07	**.77**	-.03	**.79**	.00	**.78**
8. Sales to local markets	.04	**.69**	-.04	**.73**	-.01	**.71**
1. Expansion of sales	.11	**.67**	.00	**.69**	.04	**.68**
6. Market access	.26	**.64**	.24	**.64**	.25	**.64**
% of variance explained	40.45	20.64	37.25	25.30	38.36	23.51
Confirmatory factor analysis fit indices						
Goodness of fit index (GFI)		.94		.95		.95
Adjusted GFI		.91		.91		.93
Normed fit index		.94		.93		.94
Comparative fit index		.95		.94		.95
Root mean square error of approximation		.08		.09		.08
Standardized root-mean-square residual		.05		.06		.06
Descriptives						
Scale summated means	4.26	5.22	4.25	5.15	4.25	5.18
SD	.99	.58	.93	.60	.96	.59
Scale Cronbach alpha	.89	.75	.89	.78	.89	.77
Correlation between VI and HI subscales		.29*		.14*		.20*
Covariance between VI and HI subscales accounted for by common method variance [d]		.002		-.084		-.011

(a) VI = vertical internationalization, HI = horizontal internationalization.

(b) Items are listed in descending order of vertical and then horizontal internationalization item loadings for the initial test sample. Original item order is indicated by item number.

(c) Question: Thinking generally, in deciding to have direct investments in one or more foreign economy, do you agree that the following are important motivating factors for your firm? *p<.01, **p<.001, two tailed.

(d) Difference in covariance of VI and HI between structural equation models with and without a latent factor for common method variance.

Table 3. Comparison of the VHIS between manufacturing and service MNEs

	Manufacturing MNEs N = 701		Service MNEs N = 948	
	VI subscale	HI subscale	VI subscale	HI subscale
Principal component item loadings				
11. Lower supply prices	**.88**	.07	**.89**	.08
7. Reduction of process/production costs	**.83**	.09	**.88**	.08
5. Availability of cheaper inputs	**.81**	.12	**.85**	.05
2. Lower labor prices	**.76**	-.03	**.80**	.01
9. Procurement of inputs	**.73**	.18	**.78**	.18
3. Market opportunities	.05	**.82**	.01	**.74**
10. More selling opportunities	.07	**.80**	.14	**.77**
6. Market access	.22	**.75**	.28	**.51**
8. Sales to local markets	.00	**.72**	-.01	**.72**
1. Expansion of sales	.10	**.69**	.01	**.67**
% of variance explained	38.82	24.19	38.74	21.19
Confirmatory factor analysis fit indices				
Goodness of fit index (GFI)	.95		.95	
Adjusted GFI	.92		.92	
Normed fit index	.94		.94	
Comparative fit index	.95		.95	
Root mean square error of approximation	.07		.08	
Standardized root-mean-square residual	.05		.06	
Descriptives				
Scale mean	4.44	5.15	4.11	5.20
SD	.85	.66	1.00	.54
Scale Cronbach alpha	.90	.81	.87	.71
Correlation between VI and HI subscales	.21*		.22*	
Covariance between VI and HI subscales accounted for by common method variance	.009		-.028	

*$p<.01$, **$p<.001$, two tailed.

Table 4. Country-specific application of VHIS factor analyses, reliabilities and descriptives[a]

	Full sample N = 1125		Manufacturing MNEs N = 435		Service MNEs N = 690	
	VI subscale	HI subscale	VI subscale	HI subscale	VI subscale	HI subscale
Principal component item loadings						
7. Reduction of process/production costs	**.89**	-.05	**.88**	-.06	**.90**	-.02
11. Lower supply prices	**.89**	-.05	**.88**	-.11	**.88**	.02
5. Availability of cheaper inputs	**.88**	-.05	**.84**	-.09	**.89**	.00
2. Lower labor prices	**.84**	-.05	**.79**	-.08	**.86**	-.01
9. Procurement of inputs	**.82**	.10	**.78**	.03	**.83**	.17
10. More selling opportunities	-.01	**.85**	-.04	**.87**	.04	**.82**
3. Market opportunities	-.02	**.84**	-.05	**.88**	.04	**.80**
8. Sales to local markets	-.08	**.79**	-.18	**.81**	-.03	**.78**
1. Expansion of sales	-.06	**.77**	-.08	**.79**	-.04	**.76**
6. Market access	.08	**.73**	.04	**.81**	.11	**.65**
% of variance explained	37.92	31.50	40.23	29.89	38.74	28.90
Confirmatory factor analysis fit indices						
Goodness of fit index (GFI)		.96		.96		.94
Adjusted GFI		.93		.93		.90
Normed fit index		.96		.96		.94
Comparative fit index		.97		.98		.95
Root mean square error of approximation		.07		.06		.09
Standardized root-mean-square residual		.05		.04		.06
Descriptives						
Scale summated means	4.13	5.02	4.36	4.89	3.98	5.10
SD	.99	.72	.91	.80	1.01	.64
Scale Cronbach alpha	.91	.85	.89	.89	.91	.81
Correlation between VI and HI subscales		.-.05		-.15**		.07
Covariance between VI and HI subscales accounted for by common method variance		.011		-.001		.016

(a) Question: Thinking generally, in deciding to have direct investments in China, do you agree that the following are important motivating factors for your firm?

*p<.01, **p<.001, two tailed.

Table 5. Criterion-related validity

	Correlations with VHIS subscales	
Objective variables[a]	VI subscale	HI subscale
Manufacturers[b] – general sample	.16 **	-.08 *
Manufacturers[b] – country-specific sample	.19 **	-.15 **
Financial services[b] – general sample	-.20 **	.07 *
Financial services[b] – country-specific sample	-.17 **	.15 **
Internationalization extent, sales[c] – general sample	-.07 *	.12 **
Internationalization extent, sales[c] – country-specific sample	-.09 **	.20 **
Internationalization extent, production[c] – general sample	-.05	.06
Internationalization extent, production[c] – country-specific sample	-.06	.12 **
Size[c] – general sample	-.03	.08 **
Size[c] – country-specific sample	-.01	.13 **
Subjective variables[d]		
Low cost[c] – general sample	.07 *	.05
Low cost[c] – country-specific sample	.32 **	.06
Market attractiveness[c] – general sample	-.01	.18 **
Market attractiveness[c] – country-specific sample	.02	.19 **

(a) General sample $N = 1,649$, country-specific $N = 1,125$.
(b) Point biserial correlations.
(c) Pearson product moment correlations.
(d) General sample $N = 221$, country-specific $N = 483$.
*$p<.01$, **$p<.001$, two tailed.