Lex Humana, v. 15, n.1, 2023, ISSN 2175-0947

METHOD FOR DEVELOPING A READINESS ASSESSMENT SCALE

MÉTODO PARA DESENVOLVER UMA ESCALA DE AVALIAÇÃO DA PRONTIDÃO

NGUYỄN THANH HOÀNG
University of Social Sciences and Humanities, Vietnam National University of Ho Chi Minh City
Vietnam
hoangnguyenfir@hcmussh.edu.vn

Received: 15 Sep 2022
Accepted: 15 Jan 2023
Published: 30 Jan 2023

Corresponding author:
hoangnguyenfir@hcmussh.edu.vn

Abstract: This article is to introduce a way to build a new instrument, especially a readiness assessment tool. An example case is assessing the readiness of domestic enterprises to link up with foreign enterprises in the global supply chain. The highlight of this paper is the use of available data sources - expert opinions published in the mainstream press to develop measurement criteria. The methods for checking the stability and accuracy of the instrument are introduced in detail. For studies that collect data through questionnaires, research results depend a lot on the quality of the scale. This paper contributes empirical knowledge to researchers in developing a new instrument with high reliability.

Keywords: Instrument. A scale for scientific research. Readiness level. Questionnaire.

Resumo: Este artigo visa introduzir uma forma de construir um novo instrumento, especialmente um instrumento de avaliação da prontidão. Um caso exemplar é a avaliação da prontidão das empresas nacionais para se ligarem a empresas estrangeiras na cadeia de abastecimento global. O destaque deste artigo é a utilização das fontes de dados disponíveis - pareceres de peritos publicados na grande imprensa para desenvolver critérios de medição. Os métodos para verificar a estabilidade e precisão do instrumento são introduzidos em pormenor. Para estudos que recolhem dados através de questionários, os resultados da investigação dependem muito da qualidade da escala. Este artigo contribui para o conhecimento empírico dos investigadores no desenvolvimento de um novo instrumento com elevada fiabilidade.

Palavras-chave: Instrumento. Uma escala para a investigação científica. Nível de prontidão. Questionário.

1. INTRODUCTION

Questionnaires are used in a wide range of empirical studies, from sociology to education, to medicine and economics. For studies with already existing scales (or questionnaire), researchers just borrow the scales and add some criteria suitable for the specificity and distinctiveness of the research case. For example, when a researcher wants to measure customer satisfaction with a certain service, the SERVQUAL model developed by (Parasuraman et al., 1985) is an option. Some very popular models such as the item response theory (IRT) model uses in educational achievement assessment (Lord, 2012; Rasch, 1960) model is to measure the responsive abilities, attitudes, or personality traits; technological readiness scale (TRL) model is to determine technological development levels (Héder, 2017). However, some studies do not have an available scale, researchers need to develop a new scale for their own research. Such a new scale requires stability and representativeness, which can be applied in many different cases and studies. At the same time, the scale should be a good tool to collect data for high reliability. This article has the purpose of presenting how to build a scale in scientific research. We use an empirical study to illustrate our propose. The example case is to assess the internal capabilities of local enterprises on levels corresponding to their readiness to participate in the global supply chain through linkages with foreign direct investment enterprises.

The article structure consists of six sections. The second section presents the theoretical and practical foundation for building the scale (Item Generation); the third section introduces sample selection and questionnaire distribution; the fourth section represents methods for testing the instrument. The fifth section discusses the results and finally the conclusion in the sixth section.

2. THEORETICAL AND PRACTICAL FOUNDATION FOR ITEM GENERATION

To build the scale, you need to identify (1) the major domains of the scale, then (2) the criteria to measure the factors, and (3) the relevant demographic factors (for example, if the survey subject is students, the related factors should be the year of study, major of study, gender, etc.). In this article, we also introduce (4) ways to develop tools based on different levels, namely the readiness of domestic enterprises to cooperate with foreign enterprises.

2.1 Major Domains

To build the major factors in the scale, you definitely need to base it on the theory related to your research's issue. For example, to measure customer attitudes towards a certain product, you need to use the theory of customer attitudes/behaviour which was introduced by (Ajzen, 2020; Ajzen & Fishbein, 1980; Davis, 1989); or to determine the motivations of multinational enterprises (MNEs) to invest abroad, you can use (Dunning & Lundan, 2008)'s study.

In this study, our example is to measure the internal capacity of local enterprises; therefore, the theory of firm capacity is reviewed. In other words, the main factors of the scale must be reflected the internal capacity of the enterprises. Here is how we select the major domains for our instrument to evaluate the internal capacity of local enterprises.

Enterprise competitiveness has been studied by scholars from many angles, for example, the capacity to manage the value chain from input to output (Porter, 1985); the ability to satisfy consumer needs (Market-based approach) (Baker & Sinkula, 1999; Deshpandé & Farley, 1998); the capacity to recognize the production potential of enterprises, thereby exploiting and expanding operations (Competence-based view) (Carlsson-Eliasson, 1991, quoted from (Kállay, 2012)); the capacity to identify and seize untapped market opportunities (Freiling & Verlag, 2004); management capacity of corporate managers (Eliasson, 1990a, p.238, cited from Foss, 1993). From the perspective of Resource-Based View (RBV), internal capacity was an important factor affecting business strategy of enterprises. Even within the same industry, the success of a firm depended on its internal attributes (Wernerfelt, 1995). This difference stems from differences in product quality policies, distribution channels, competitive positioning, and financial leverage (Foss, 1993). It is these firm-specific factors that allow companies to gain a sustainable competitive advantage, which translates into high performance in the marketplace (Barney, 1991, 1997; Day, 1994; Wernerfelt, 1984; cited from (Lee & Rugman, 2012)). Whether standing from different angles looking at the competitiveness of enterprises; However, all schools have the common view that internal resources of enterprises are the root of the existence and effectiveness of enterprises in maintaining competitive advantages (Freiling & Verlag, 2004). Compiled from many studies, (Islami et al., 2018) recorded the basic internal competencies of an enterprise that were the quality of the workforce including knowledge and skills which accumulated through company training (Song et al., 2003), experience (Hoffman et al., 1998), ability to work in groups (Cooper, 1990); the level of technology shown through research

and development activities (Bhattacharya & Bloch, 2004); and corporate governance capacity (Webster, 2004). (Ahmedova, 2015) also synthesized from previous studies five groups of factors affecting the competitiveness of enterprises, including finance, innovation activities, intellectual property, national internationalisation and implementation with optimal solutions.

In short, the internal capacity of an enterprise can be summarised by covering the quality of human resources, technology level, corporate governance, supply as well as financial capacity. These are the five major theoretical factors included in the framework.

2.2 Indicators

There are several ways to define measurement criteria. You can refer to the observed variables that previous studies have used, or get invited experts' opinions. Summarizing articles on scale development, (Morgado et al., 2017) reported that 74.2% (on a sample of 105 articles) used experts' opinions for designing measurement criteria. In this study, the factors used to measure the five main factors of the research framework are built from expert opinions on the issue of linking domestic enterprises with FDI enterprises. Instead of gathering opinions from a closed source by an invited group of specific experts, this research collected opinions from an open-source newspaper. One precedent is the work of (Pandit, 1996) when he used documents in the form of newspaper reports, trade journals, business journals, government publications, broker reviews, public documents annual company reports, and press releases to develop the theory of corporate turnaround.

To determine the measurement elements for the five main factors of the instrument, we collected opinions of experts who were knowledgeable about linkages between domestic and FDI enterprises in Vietnam. These opinions were published in 68 articles in mainstream journals to ensure legitimacy. Collected articles were posted from 2016 to 2022 to ensure objectivity over time. With this open data source, we collected multi-dimensional data (multi-sectors and multi-local) to ensure objectivity and reality to include in the scale. The experts were policy makers and managers of domestic and FDI companies; therefore, they were trustworthy knowledgeable people. Table 1 describes an example of extracting expert opinion into scale's indicators.

Table 1. Sample for extracting expert's opinion into measurement criteria

Newspaper	Expert's opinion	Indicators
Small businesses join the global supply chain	*Most of Vietnam's small and medium enterprises (SMEs) have limited human*	- Weakness of internal capacity in terms of

Posted on Nhandan[1] (Tô Hà, 2020)	*resources, including low levels of corporate governance. It is difficult for them to accumulate and concentrate capital to invest in technological innovation, expanding production scale, and improving productivity and product quality. Therefore, Vietnamese enterprises have not been able to meet the increasingly strict requirements of major domestic and international partners.* *In addition, the business leadership team has a cautious mentality, and low risk-taking ability in investing in upgrading supply standards; thus, they can't breakthrough yet.* Speech by Mr. Nguyen Chi Dung, Minister of Planning and Investment of Vietnam.	human resource, technology, finance, production, product quality, supply ability, and governance corporation. -Awareness of international linkage.

The first draft had 24 items to measure five dimensions of internal capacity in Corporate Governance (denoted as G), Human Resources (denoted as H), Finance (denoted as F), Technology (denoted as T), and Supply (denoted as S); and 3 items to present the Perception & Readiness for Linkages (denoted as L). Table 2 in Section 5.1 presents these measurement elements.

2.3 Demographic Factors

Due to the sensitivity of enterprises when asked questions related to their business; the proposed demographic factors were mainly collected on the business field, year of establishment, and size of the firm. Besides, the working experience of respondents was also paid attention to ensure the reliability of responses.

For the business field: the linkage between local and FDI enterprises can take place horizontally and/or vertically. In this study, the selected local enterprises must work in the same industry that the majority of FDI enterprises invest in or/and in the industries that can become a supplier or distributors of FDI enterprises. Our subsectors include agriculture, forestry and fisheries; processing and manufacturing industries; production and distribution of electricity and water; and other industries such as construction, wholesale and services.

For the establishment years: Referring to the research of (Fariñas & Moreno 2000), and (Loderer and Waelchli 2010), the authors classified enterprises by 5 years incrementally. In

[1] Nhan Dan Newspaper, Central Authority of the Communist Party of Vietnam, Voice of the Party, State and People of Vietnam. https://nhandan.vn/sp/AboutUs

this study, we also apply the same. In addition, only firms with a minimum operating period of one year are included in the study.

In terms of business size: To identify the size of enterprises, we based on Article 6 of Decree 39/2018 / ND-CP of Vietnam. Enterprise size in Vietnam is divided into 4 levels: micro, small, medium and large. There are 3 criteria to determine the size of an enterprise, including: the average number of employees participating in social insurance each year; total revenue of the year; or total capital of the year. These criteria have differences between two groups of fields: agriculture, forestry and fishery; and industrial and construction or commercial and service sectors.

Regarding respondents' experience: This information is to ensure the reliability of respondents' opinions. The higher the manager position, the higher the full awareness of the company's activities. In our study, statistics show that 48.10% of respondents hold a managerial position for 5 years or more in the business.

2.4 Scale for Readiness Level

To build a multi-level functional scale, we refer to previous studies on building similar scales. Several research models[2] used a level of readiness. Readiness is understood as a state of being ready (Merriam-Webster); or willing to do something (Collins Dictionary). The scale is a continuum ranging from low-level to high-level readiness.

For application in economics, one of the known, widely used and uniform scales was the technological readiness scale (TRL) first developed by NASA and adopted by the European Union (Héder, 2017). TRL was used to evaluate the maturity of technology on a scale of 0-9: Level 1 was fundamental and Level 9 was the use of modern technology. For the same purpose, (Lichtblau et al., 2014) proposed the 5-point Likert scale with 6 levels: Level 0: "Outsider", Level 1: "Beginner", Level 2: "Intermediate", Level 3: "Experienced", Level 4: "Expert", Level 5: "Top performer". Developing from existing TRL models, (Akdil et al., 2018) suggested a model for measuring the technological maturity of enterprises (Industry 4.0 maturity model) with four levels, Level 0: "Absence", Level 1: "Existence", Level 2: "Survival", and Level 3: "Maturity". To determine the "maturity" level, questions were weighted from 0 to 3 corresponding to each level. In the field of management of

[2] Technology Readiness Level (TRL), Investment Readiness Level (IRL), Market Readiness Level (MRL), Manufacturing Readiness Level (MRL), Business Readiness Level (TRL), Innovation Readiness Level (TRL), Capacity Maturity Model (CMM).

organizational behaviour, (Novit et al., 1971) built up a model that measured the capacity and the willingness of individuals to perform tasks. This scale ranged from low, low to medium, medium to high, and high.

To some extent, the capacity of a local firm will correspond to the level of willingness to join the global supply chain. In this study, the tool is an interval scale with 5 levels named according to the level of the business' scope of activities[3]:

- Level 1 "Local" (corresponding to 0 - 1 point, or 0-25%): business capacity is very low, small markets, not willing to change their business strategies.

- Level 2 "Regional" (corresponding to 1 - 2 points, or 26-50%): business capacity is quite low, small scale, low technology; low readiness to join the global supply chain.

- Level 3 "National" (corresponding to 2 - 3 points, or 51-75%): enterprise capacity at fair level, medium to low level of large scale; or having resources belonging to the VRIN group and seeking international markets.

- Level 4 "International" (corresponding to 3 - 4 points, or 76-90%): strong capacity, large scale; modern technology; high readiness to join the global supply chain.

- Level 5 "Global" (corresponding to 4-5 points, or 91-100%): powerful ability, large scale; modern technology; leading the market; very high readiness or already joined the global supply chain.

3. SAMPLE SELECTION AND QUESTIONNAIRE DISTRIBUTION

We award that the quality and quantity of the sample will significantly affect the quality of the tool. To ensure the generalizability of the new scale, the sample needs to be selected in a variety of compositions (J. C. Nunnally, 1967).

In terms of the sample nature, companies in our sample were selected according to the following criteria: 1) at least one year of operation; 2) diversifying the business lines; 3) diversifying the size of capital and labour including all medium and large enterprises while small companies were selected randomly.

Regarding the sample size, in the case of developing a new scale, (DeVellis, 2016) suggested a sample size of 300 for a unidimensional scale constructed with about 20 criteria. Besides,

[3] By chance, we found a coincidental idea on a website writing about stages evolution of business, posted by Akrani on September 22, 2012, at https://kalyan-city.blogspot.com/2012/09/stages-in-evolution-of-business-from.html

in a model using the CFA test, with 3 to 4 observed variables for each factor, the sample size must be greater than 100 (Boomsma, 1985; Marsh and Hau, 1999). To ensure sample power, (Gudgeon et al., 1994) divided the sample into six levels: very poor: 50, poor: 100, fair: 200, good: 300, very good: 500 and excellent: 1000. Similarly, (Tabachnick & Fidell, 2012) recommended a sample of around 300. With the rule of thumb, the sample size can be determined based on the ratio of sample size (N) and several measurement criteria (p). The widely accepted ratio is 10: 1 (J. Nunnally & Bernstein, 1994). Based on the indexes just listed, our final tool has six dimensions with 25 indicators, averaging 4 variables per dimension (factor); the sample of 300 was sufficient.

Our empirical study is conducted in the Ben Tre province of Vietnam. According to the Department of Planning and Investment of Ben Tre Province (DPI), there are 824 enterprises belonging to the above-defined sample group. With 95% confidence and a 5% margin of error, the sample should be 280; if the confidence level is up to 99%, the sample size is 363 (Krejcie & Morgan, 1970).

The quality of the questionnaire was strictly controlled. Unsatisfactory responses were removed and replaced by sending questionnaires to other businesses. The research team stopped at accepting 308 copies, reaching the rate of 102.66% of the expected number of samples, and getting more than one-third of the population. This confirmed the sample can contribute to the generalizability of the new scale.

4. METHODS FOR TESTING THE INSTRUMENT

In scientific research, testing the accuracy and consistency of a scale is very important. This study aims to develop a new scale based on a questionnaire; therefore, the quantitative tests for the questionnaire is of interest.

According to (Golafshani 2015), reliability was the degree to which the results did not change over time, represented the study as a whole, and applied similarly. Reliability testing was to test the stability and consistency of the scale (Carmines and Zeller 1979; Alheide and Johnson 1998). Meanwhile, testing the accuracy (Validity) to ensure the truthfulness (Mehrens and Lehmann 1987) and completeness of the scale. (Field 2013) argues that validity was "measuring what a study intends to measure". (Alheide and Johnson 1998) concluded that reliability referred to the consistency of the results, while validity denoted the truthfulness of the results.

Various testing techniques have been developed that easily find in some articles or review articles. Though there are differences in subgroups of each test type; however, the nature and use of the testing techniques are relatively consistent. In this paper, we shortly represented a summary of the (Bolarinwa 2015) techniques used. This article synthesizes the tests for questionnaires.

4.1 Reliability Testing

Stability and consistency are tested through the following techniques:

(1) Repeatability / Test-retest reliability: this test reflects a similar level of score test between two times of the test on the same objects and the same method.

(2) Alternative / Parallel-form Reliability or Equivalence: this test reflects a correlation of two forms that are not exactly similar on the same scale (same questionnaire but changing order questions and wording) conducted on the same subject. The higher the correlation, the more equivalent the two questionnaires. This method was rarely used because it lacked a basis to validate that these were two parallel tests.

(3) Inter-rater reliability / Inter-observer reliability: this test reflects similar testing results among independent reviewers who conducted the same questionnaire.

(4) Internal consistency / inter-item consistency / or Homogeneity: this test evaluates the consistency of items on the scale; the scale has high internal consistency when the measured variables on the scale are closely related.

Internal consistency is estimated through the Split-half reliability index, the coefficient alpha index, and the Kuder-Richardson formula 20 (KR-20) index.

- *Split-half reliability index* measuring the correlation between two questionnaires split from a common questionnaire, answered by the same respondents.

- *Coefficient alpha index* reflects the degree of a close correlation between observed variables in the same factor, these observed variables express the characteristics of that factor. Most of the studies used this index. Many authors used different cut-offs for different studies.

Alpha coefficients are often used for criteria with several choices (e.g., Likert-5, Likert-7); while KR-20 is used for dichotomous variables.

4.2 Validity Testing

Tests of validity are divided into two subgroups: theoretical construct validity and empirical construct validity.

(1) The theoretical construct validity: this test is to evaluate how well constructs[4] from the theory are performed in a scale (questionnaire). Two tests can be done alone or together:

- *Face validity* shows the level of the experts' assessment that the observed variables of the scale have fully measured the characteristics and traits of the research issues.

- *Content validity* shows the level of experts' assessment that the observed variables are representative enough, and cover all research aspects and the other points such as readability, comprehension, and clarity.

There are two ways to evaluate content validity. The first way is in a binary form likes "Favorable" or "Unfavourable". The second way commonly used today is on a scale (for example from 1 through 4, or 1 to 5). There are two quantities that have to calculate[5]: I-CVI (the Item Content Validity Index) and S-CVI (the whole Scale Content Validity Index).

(2) Empirical construct validity:

The validity tests are divided into two subgroups: Criterion validity and Construct validity:

- Criterion validity / Reference validity: the test is to compare the correlation coefficiency between the new scale and another accepted measurement standard in the same research issue.

There are two pieces of evidence to assess:

+ *Concurrence* measuring the correlation between the results of the new scale and a currently used standard.

+ *Predictive* measuring the degree of correlation in which the new scale can be used to measure the research issue and its results will be used for predicting another issue in the future.

- Construct validity: this test is to measure whether the new scale covers all characteristics or corresponds to the theoretical framework of the research issue. There are five types of evidence used according to the study's purposes:

+ *Convergent validity* measures whether applying the same concept in different ways (for example, using a questionnaire and doing observations) with the same objects, but giving similar results.

[4] Crocker and Algina (1986) called "construct" as "informed scientific imagination"; Nunnally and Bernstein (1994) defined a construct as a hypothesis.

[5] Read Polit et al. (2007) Is the CVI an Acceptable Indicator of Content Validity? Appraisal and Recommendations

+ *Discriminant validity* measuring the difference between one concept and another related concept when using the same measurement.

+ *Known-group validity* shows the difference between a "known group" which have already had knowledge or experience about the research issue and an "unknown group"; in which, the "known group" expects to give higher measurement scores than the other group.

+ *Factorial validity*: in a scale involving several factors (dimensions), in each factor there are many observed variables (items), the factorial validity provides evidence of a correlation between the variables in the same factor instead of variables in other factors. The techniques used are Exploratory Factor Analysis (EFA) and Confirmation Factor Analysis (CFA).

+ *Hypothesis-testing validity* provides evidence to support the research hypothesis about the relationship between the factors which is built from the theoretical basis.

As mentioned, depending on the research's purposes, the scale will be tested by respective techniques. Based to the (Morgado et al. 2017), most of the authors used either EFA (88.6%), CFA (72.3%), convergent validity (72.3%), or discriminant validity (56.2%) for testing the construct validity; while 65.7% of the research applied both EFA and CFA. To test the reliability of the scale, internal consistency is used by most authors; while 22.8% of research studies use the extra test-retest reliability.

In this study, the tool is tested in corresponding steps with the following techniques:

Step 1. Testing the theoretical construct validity through a combined assessment of Face validity and Content validity.

Step 2. Testing the reliability under Internal consistency by comparing the coefficient alpha indices of the whole sample and its sub-groups.

Step 3. Testing the empirical construct validity in the form of Construct validity by using Factorial validity including EFA and CFA. Details were presented in the Results section.

Collected data were analyzed by SPSS and AMOS.

5. DATA ANALYSIS AND RESULT

5.1 Face and Content Validity

The first draft of the questionnaire is submitted for assessment by an expert panel (EP) including two researchers and two local governmental administrators. The evaluation focuses

on "the appropriateness of the observed variables measuring the main factor" and "the level of comprehension and clarity" of the words used in the questionnaire. The rating includes four points: 1 = very inconsistent / very unclear to 4 = very suitable, very clear. In the face-to-face meeting, EP takes turns evaluating each observed variable, and the inconsistent and unclear points are adjusted during the meeting. The EP group proposes to plus two additional items measuring the level of "proactively seeking linkages with FDI" and "joining professional associations". Factor Perception & Readiness for Linkages (L) has 5 items.

The second draft is sent to a target panel (TP) of six local CEOs representing 3 sizes of businesses: small, medium, and large. Similar to EP, the TP meet online and evaluates each observed variable. The TP proposes to gather items to condense the number of questions so that businesses can easily make choices. Specifically, grouping two items "Access to bank loans" and "Ability to expand business capital" into "Access to capital"; abolishes the item "The need to raise capital" because there is a criterion "Capital absorbability". Two items measuring the experience and education of managers are grouped into one item; likewise, for the criteria measuring the capacity of employees. Factor Finance (F) has 3 items, and factor Human Resources (H) has 3 items.

The third draft is re-sent to the EP for grading. Most experts rate 3 and 4 for the items, except 5 items rate at 2: G2 (TP5), H2 (EP1), F2 (EP3), T4 (TP1) and L5 (TP2).

The accuracy of the content value is assessed by the I-CVI and S-CVI indexes.

The I-CVI is calculated on each observed variable, which is the ratio between the number of experts who agree (choose level 3 and level 4) to the total number of experts participating in the assessment (Polit et al., 2007). The variables G2, H2, F2, T4, and L5 have 9 out of 10 experts agree, the I-CVI value of these variables is 0.9. The remaining variables had 10/10 experts agree, that the index I-CVI is 1. The value of acceptance of the I-CVI depends on the number of experts. If the number is five or more, this indicator is at least 0.83.

The S-CVI index is calculated by the average of all I-CVI in the scale (Lynn, 1986, cited Polit *et al.*, 2007). The value of S-CVI is 0.98 (((20x1)+(5x0.9))/25), which is greater than the standard value of 0.8 (Davis 1992, cited Polit *et al.*, 2007). This shows that the items in this scale satisfy the content validity.

Thus, the final scale (Table 2) represents a scale of 5 levels from low to high: level 1 = very low capacity / low readiness to level 5 = perfect capacity / very high readiness, with 25 indicators to measure six domains of Corporate Governance (5 items: G1-G5), Human

Resources (3 items: H1-H3), Finance (3 items: F1-F3), Technology (5 items: T1-T5), Supply (4 items: S1-S4) and Perception & Readiness for linkages (5 items: L1-L5).

Table 2. The instrument for measuring readiness levels in terms of the company's internal capacity

Factors	Symbol	Indicators
CORPORATE GOVERNANCE	G1	Daily executive management
	G2	PR management
	G3	Production and service management
	G4	Supply chain management
	G5	Strategic management
HUMAN RESOURCES	H1	Quality of labours/staff: Knowledge, Skills and Experience
	H2	Quality of management team: Knowledge, Skills and Experience
	H3	Ability to participate in association activities in terms of HR
FINANCE	F1	Ability to accumulate capital and equity capital
	F2	Ability to access bank loans and expand business capital
	F3	Capital absorption (deployment and use)
TECHNOLOGY	T1	Application of information technology to the management and administration of business activities
	T2	Application of technology in production / cultivation / animal husbandry / service
	T3	Application of technology in supply chain
	T4	R&D
	T5	Absorption of technology
SUPPLY	S1	Quality and Price
	S2	Quantity
	S3	Diversity of products and services
	S4	Experience
PERCEPTION & READINESS FOR LINKAGE	L1	Awareness of the importance of linking with FDI
	L2	Potential ability to linkages
	L3	Current linkages
	L4	Readiness to linkages
	L5	Member of professional associations

5.2 Internal Consistency

The scale is assessed for reliability through internal consistency evaluation. As stated, we divide the sample into two sub-groups: group 1 "2015 onwards" with an active life of 5 years or less, and group 2 "2014 backwards" with a history of more than 5 years. The difference in the firm age might partly reflect the difference in the firm's capacity. If the measurement scores between the three groups do not have a great distance, the scale is stable.

Many studies use Cronbach's alpha coefficients for confirming the degree of internal consistency. However, the alpha coefficient does not have a common standard for determining uniformity. (Cortina, 1993) asserted that this coefficient depended on the number of items/indicators or the questions on the scale. The scale with more items had a higher coefficient. Compiling from studies using Cronbach's alpha coefficient, (Peterson, 1994) identified a typical average of 0.77, where the scale was considered homogeneous. (Hoang Trong & Chu Nguyen Mong Ngoc, 2008) classified a scale as very good if the CA coefficient was from 0.8 to close to 1; good scale: 0.7 - 0.8; acceptable scale: from 0.6 and up. Notably, (Hulin et al., 2001) suggested that a value higher than 0.95 may be an indication of redundancy. Moreover, according to (J. Nunnally & Bernstein, 1994) only observed variables with correlation coefficients greater than 0.3 can be retained; if below 0.3, they will be discarded.

In this study, after performing EFA, there are two excluded observations, G2 and H2. The results in Table 3 show that the CA coefficient of variables was from 0.601 to 0.862; there are no observed variables with correlation coefficients less than 0.3. Furthermore, there is no big difference in the homogeneity of the scales between the three groups. This confirms that the scale is stable and consistent.

Table 3. Internal consistency of the scale

	Items	2015 onwards (n=144)	2014 backwards (n=164)	Overall (n=308)
Corporate Governance	5-1	0.797	0.862	0.835
Human resources	3-1	0.798	0.852	0.831
Finance	3	0.685	0.601	0.639
Technology	5	0.776	0.677	0.722
Supply	4	0.707	0.659	0.680
Perception & Readiness for linkages	5	0.689	0.831	0.713

5.3 Factorial Validity

The test the accuracy of the scale to ensure the correctness and completeness of the selected criteria or the observed variables.

At first, we apply EFA. Acceptable measurement variables include a KMO coefficient of about $0.5 \leq KMO \leq 1$ and the Bartlett test with Sig. <0.05 (Hoang Trong & Chu Nguyen Mong Ngoc, 2008); the loading factor must be greater than 0.4 to ensure stability (Guadagnoli and Velicer, 1988, cited Costello & Osborne, 2005).

Factor analysis is not simple because of the variety of extractions and rotations. According to (UCLA: Statistical Consulting Group., 2018), there are three main extraction methods: Principal components analysis (PCA), Principal axis factoring (PAF) and Maximum Likelihood. Besides that, there are two main rotations: Orthogonal rotation with the common use of Varimax and Oblique rotation with the common use of Promax, then Direct Oblimin.

Factor analysis helps the dataset have a simple structure. PAF is useful if handling samples in preparation for further analysis; while Maximum Likelihood or Kaiser's alpha factoring should be applied to the new scale development (Field, 2013: 674-675). Table 4 below shows the dataset applying different extraction and rotation methods. The loading factors are not too different, showing that the scale is stable.

Table 4. A comparison of factor loading between PCA-PAF-Alpha factoring and CFA

Extraction Method		PCA	PAF	Alpha Factoring	CFA
Rotation Method		Orthogonal	Oblique	Oblique	
		Varimax	**Promax**	**Promax**	
Item Loadings		Factor Loading			SLE
F_G	G1	0.86	0.95	0.90	0.93
	G3	0.82	0.88	0.86	0.88
	G4	0.66	0.45	0.55	0.57
	G5	0.68	0.47	0.57	0.61
F_T	T1	0.62	0.64	0.54	0.58
	T2	0.69	0.69	0.59	0.54
	T3	0.69	0.59	0.63	0.65
	T4	0.64	0.64	0.54	0.57
	T5	0.64	0.69	0.60	0.59
F_S	S1	0.56	0.50	0.47	0.69
	S2	0.70	0.66	0.67	0.52
	S3	0.65	0.57	0.56	0.62

	S4	0.65	0.53	0.54	0.52
F-F	F1	0.63	0.43	0.44	0.50
	F2	0.67	0.61	0.66	0.81
	F3	0.77	0.62	0.71	0.54
F_H	H1	0.85	0.77	0.82	0.64
	H3	0.87	0.78	0.72	0.89
F_L	L1	0.71	0.79	0.80	0.83
	L2	0.77	0.87	0.85	0.81
	L3	0.59	0.40	Freestanding	Suppressed
	L4	0.72	0.71	0.61	0.69
	L5	0.59	0.42	Freestanding	Suppressed

After eliminating two variables L3 and L5 respectively, KMO is 0.814; Bartlett's equals 0,000; at the eigenvalues of 1.061 is extracted six factors with total variance explained reaching 62,785%. As such, all of these criteria qualify for inclusion in the next analysis.

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.814
Bartlett's Test of Sphericity	Approx. Chi-Square	2223.740
	df	210
	Sig.	.000

(Costello & Osborne, 2005) argued that "Once an instrument has been developed using EFA and other techniques, it is time to move to confirmatory factor analysis". (Hair et al., 2010) believe that factor loading estimates are accepted from 0.5, preferably 0.7 and over. Standardized Loading Estimates (SLE) in Table 4 are from 0.5 to 0.9. This result confirms that the scale is valid.

(Bolarinwa, 2015) mentioned that some authors used hypothesis testing in evaluating the accuracy of the new scale. In this study, the regression result for R^2 is 0.414; F is 42.749 and Sig is equivalent to 0.000. The result supports our instrument.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.644[a]	.414	.405	1.68826

a. Predictors: (Constant), F_H, F_F, F_T, F_G, F_S

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	609.223	5	121.845	42.749	.000[b]
	Residual	860.764	302	2.850		
	Total	1469.987	307			

a. Dependent Variable: F_L
b. Predictors: (Constant), F_H, F_F, F_T, F_G, F_S

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.289	.919		1.404	.161
	F_G	.332	.046	.370	7.272	.000
	F_T	.241	.042	.269	5.766	.000
	F_S	.103	.055	.099	1.895	.059
	F_F	.186	.057	.161	3.275	.001
	F_H	.075	.053	.063	1.420	.157

a. Dependent Variable: F_L

Generally, the results of testing Face and Content validity, Internal consistency and Factorial Validity determined that the scale met the conditions of reliability and validity and is ready for use.

6. CONCLUSION

Building a new scale requires a lot of work.

First of all, you have to rely on an appropriate theoretical foundation to build a research framework that includes major domains. Observable variables play a very important role when they are used to measure the key factors. In a new scale, these items are distilled from expert opinions or beneficiaries. In this paper, we propose an available and reliable source, expert opinion published in news, where experts have to filter their opinions before publicizing. In addition, the appropriateness of the width, nature of the sample, and the method of collecting information will contribute to the reliability of the data.

At the same time, the full and thorough implementation of the necessary tests is to eliminate confounding variables and keep only relevant variables. The test results are a "measurement" for the stability and accuracy of this new scale. Tests are often time-consuming and must be performed many times until a reliable and universal scale is obtained. Scale builders must be extremely patient.

Limitation: The scale's stability and accuracy will be evaluated based on inputs that are collected via respondents' choices (questionnaire). A very good scale under experts' judgment (high content validity) can become a bad scale (low CA, KMO, SLE) if the respondents are not attentive and accountable for their choices. People in managerial

positions are generally not interested in scientific research. Consequently, studies in which managers are participants can give biased results. Therefore, the excluded criteria in this study may be useful in other studies as well.

REFERENCES

Ahmedova, S. (2015). Factors for Increasing the Competitiveness of Small and Medium-Sized Enterprises (SMEs) in Bulgaria. *Procedia - Social and Behavioral Sciences*, *195*, 1104–1112. https://doi.org/10.1016/j.sbspro.2015.06.155

Ajzen, I. (2020). The theory of planned behavior: Frequently asked questions. *Human Behavior and Emerging Technologies*, *2*(4). https://doi.org/10.1002/hbe2.195

Ajzen, I., & Fishbein, M. (1980). Understanding attitudes and predicting social behaviour. New Jersey: Prentice-Hall. *Englewood Cliffs*, *50*.

Akdil, K. Y., Ustundag, A., & Cevikcan, E. (2018). *Maturity and Readiness Model for Industry 4.0 Strategy*. https://doi.org/10.1007/978-3-319-57870-5_4

Baker, W., & Sinkula, J. (1999). Learning Orientation, Market Orientation, and Innovation: Integrating and Extending Models of Organizational Performance. *Journal of Market - Focused Management*, *4*(4). https://doi.org/10.1023/A:1009830402395

Bhattacharya, M., & Bloch, H. (2004). Determinants of innovation. *Small Business Economics*, *22*(2). https://doi.org/10.1023/B:SBEJ.0000014453.94445.de

Bolarinwa, O. A. (2015). Principles and methods of validity and reliability testing of questionnaires used in social and health science researches. *Nigerian Postgraduate Medical Journal*, *22*(4), 195. https://doi.org/10.4103/1117-1936.173959

Boomsma, A. (1985). Nonconvergence, improper solutions, and starting values in lisrel maximum likelihood estimation. *Psychometrika*, *50*(2). https://doi.org/10.1007/BF02294248

Cooper, R. G. (1990). New Products: What Distinguishes the Winners? *Research-Technology Management*, *33*(6). https://doi.org/10.1080/08956308.1990.11670699

Cortina, J. M. (1993). What Is Coefficient Alpha? An Examination of Theory and Applications. *Journal of Applied Psychology*, *78*(1). https://doi.org/10.1037/0021-9010.78.1.98

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research and Evaluation*, *10*(7). https://doi.org/10.7275/jyj1-4868

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly: Management Information Systems*, *13*(3). https://doi.org/10.2307/249008

Deshpandé, R., & Farley, J. (1998). Measuring Market Orientation: Generalization and

Synthesis. *Journal of Market-Focused Management*, *2*(3). https://doi.org/10.1023/A:1009719615327

DeVellis, R. F. (2016). Scale Development Theory and Applications (Fourth Edition). *SAGE Publication*, *4*.

Dunning, J. H., & Lundan, S. M. (2008). Institutions and the OLI paradigm of the multinational enterprise. *Asia Pacific Journal of Management*, *25*(4). https://doi.org/10.1007/s10490-007-9074-z

Field, A. (2013). Discovering statistics using IBM SPSS statistics. In *Statistics* (Vol. 58).

Foss, N. J. (1993). Theories of the firm: contractual and competence perspectives. *Journal of Evolutionary Economics*, *3*(2), 127–144. https://doi.org/10.1007/BF01213830

Freiling, J., & Verlag, R. H. (2004). A competence-based theory of the firm. *Management Revue*, *15*(1), 27–52.

Gudgeon, A. C., Comrey, A. L., & Lee, H. B. (1994). A First Course in Factor Analysis. *The Statistician*, *43*(2). https://doi.org/10.2307/2348352

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). Multivariate Data Analysis. In *Vectors*. https://doi.org/10.1016/j.ijpharm.2011.02.019

Héder, M. (2017). From NASA to EU: The evolution of the TRL scale in Public Sector Innovation. *Innovation Journal*, *22*(2).

Hoang Trong, & Chu Nguyen Mong Ngoc. (2008). *Analyzing research data with SPSS, Part 1 and Part 2*. Hong Duc.

Hoffman, K., Parejo, M., Bessant, J., & Perren, L. (1998). Small firms, R&D, technology and innovation in the UK: A literature review. *Technovation*, *18*(1). https://doi.org/10.1016/S0166-4972(97)00102-8

Hulin, C., Netemeyer, R., & Cudeck, R. (2001). Can a Reliability Coefficient Be Too High? *Journal of Consumer Psychology*, *10*(1).

Islami, X., Mulolli, E., & Skenderi, N. (2018). The Effect of Factors Industrial and Internal to the Firm's Performance. *Acta Universitatis Danubius. Œconomica*, *14*(5).

Kállay, B. (2012). Evolutionary and competence based theory. *Journal of International Studies*, *5*(1), 38–45. http://www.geoffrey-hodgson.info/user/image/evcomp.pdf

Krejcie, R. V., & Morgan, D. W. (1970). Determining Sample Size for Research Activities. *Educational and Psychological Measurement*, *30*(3), 607–610.

Lee, I. H., & Rugman, A. M. (2012). Firm-specific advantages, inward FDI origins, and performance of multinational enterprises. *Journal of International Management*, *18*(2).

https://doi.org/10.1016/j.intman.2011.11.001

Lichtblau, K., Stich, V., Bertenrath, R., Blum, M., Bleider, M., Millack, A., Schmitt, K., Schmitz, E., & Schröter, M. (2014). Impuls - Industrie 4.0 Readiness. *Impuls-Stiftung Des VDMA*.

Lord, F. M. (2012). Applications of item response theory to practical testing problems. In *Applications of Item Response Theory To Practical Testing Problems*.

https://doi.org/10.4324/9780203056615

Morgado, F. F. R., Meireles, J. F. F., Neves, C. M., Amaral, A. C. S., & Ferreira, M. E. C. (2017). Scale development: Ten main limitations and recommendations to improve future research practices. *Psicologia: Reflexao e Critica*, *30*(1), 1–20. https://doi.org/10.1186/s41155-016-0057-1

Novit, M. S., Hersey, P., & Blanchard, K. H. (1971). Management of Organizational Behavior: Utilizing Human Resources. *Industrial and Labor Relations Review*, *24*(3). https://doi.org/10.2307/2521543

Nunnally, J., & Bernstein, I. (1994). Psychometric Theory, 3rd edn, 1994. *McGraw-Hill, New York*, *3*.

Nunnally, J. C. (1967). *Psychometric Theory*. McGraw-Hill.

Pandit, N. R. (1996). The Creation of Theory: A Recent Application of the Grounded Theory Method. *The Qualitative Report*, *2*(4), 1–15. https://doi.org/10.46743/2160-3715/1996.2054

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1985). A Conceptual Model of Service Quality and Its Implications for Future Research. *Journal of Marketing*, *49*(4). https://doi.org/10.1177/002224298504900403

Peterson, R. A. (1994). A Meta-Analysis of Cronbach's Coefficient Alpha. *Journal of Consumer Research*, *21*(2). https://doi.org/10.1086/209405

Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Focus on research methods: Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing and Health*, *30*(4). https://doi.org/10.1002/nur.20199

Porter, M. (1985). Competitive Advantage: Creating and sustaining superior performance. *The Free*.

Rasch, G. (1960). Studies in mathematical psychology: I. Probabilistic models for some intelligence and attainment tests. In *Studies in mathematical psychology: I. Probabilistic models for some intelligence and attainment tests.*

Song, J., Almeida, P., & Wu, G. (2003). Learning-by-hiring: When is mobility more likely to facilitate interfirm knowledge transfer? *Management Science*, *49*(4). https://doi.org/10.1287/mnsc.49.4.351.14429

Tabachnick, B. G., & Fidell, L. S. (2012). Using multivariate statistics (6th ed.). In *New York: Harper and Row*.

Tô Hà. (2020, October 26). Doanh nghiệp nhỏ tham gia chuỗi cung ứng toàn cầu. *Nhandan*. https://nhandan.vn/chuyen-lam-an/doanh-nghiep-nho-tham-gia-chuoi-cung-ung-toan-cau-621892/

UCLA: Statistical Consulting Group. (2018). PRINCIPAL COMPONENTS (PCA) AND EXPLORATORY FACTOR ANALYSIS (EFA) WITH SPSS Overview. *UCLA Institute for Digital Research & Education Statistical Consulting*, 1–43. https://stats.idre.ucla.edu/spss/seminars/efa-spss/

Webster, E. (2004). Firms' decisions to innovate and innovation routines. *Economics of Innovation and New Technology*, *13*(8). https://doi.org/10.1080/1043859041000168682 4

Wernerfelt, B. (1995). The resource-based view of the firm: Ten years after. *Strategic Management Journal*, *16*(3). https://doi.org/10.1002/smj.4250160303