VALIDATION OF SUSTAINABLE DEVELOPMENT PRACTICES SCALE USING THE BAYESIAN APPROACH TO ITEM RESPONSE THEORY

VREDNOVANJE LJESTVICE PRAKSI ODRŽIVOG RAZVOJA KORIŠTENJEM BAYESOVOG PRISTUPA U TEORIJI ODGOVORA NA ZADATKE

UDK 65.017.3:502.131.1]:519.22
Pregledni rad
Review

Martín Hernani Merino, Ph. D.
Research Professor
Universidad del Pacífico
Av. Salaverry 2020, Jesús María, Lima, PERU
Phone: ++51 1 219 0100; Mobile: ++51 975 126 669
Fax: ++51 1 470 9727
E-mail: mn.hernanim@up.edu.pe

Enver Gerald Tarazona Vargas, M. Sc.
Professor
Pontifical Catholic University of Peru (PUCP)
Av. Universitaria 1801, San Miguel, Lima 32, PERU
Phone: ++51 1 626 2000
E-mail: enver.tarazona@pucp.pe

Antonieta Hamann Pastorino, Ph. D.
Professor
Economics and Administration School - Esan University
Alonso de Molina 1652 Monterrico, Santiago de Surco – Lima 33, PERU
Phone: ++51 1 479 1987; Mobile: ++51 99 926 551
Fax: ++51 1 479 1987
E-mail: ahamann@esan.edu.pe

José Afonso Mazzon, Ph. D.
Full Professor
Administration and Economics School, University of São Paulo
Av. Luciano Gualberto, 908 - sala E187 - São Paulo, BRAZIL
Phone: ++55 11 3091 5995; Mobile: ++55 11 99688 1816
Fax: ++55 11 3818 4026
E-mail: jamazzon@usp.br

Ključne riječi:
mikro i mala poduzeća, održivi razvoj, teorija odgovora na zadatke

Key words:
micro and small enterprise (MSE), sustainable development, Item Response Theory (IRT)

SAŽETAK

Važnost kreiranja mjernih instrumenata za mjerenje učinkovitosti ekonomskog, društvenog

ABSTRACT

There has been growing recognition of the importance of creating performance measure-

i upravljanja okolišem mikro i malih poduzeća sve je priznatija. U tom kontekstu cilj ovog istraživanja jest vrednovanje instrumenta za procjenu percepcija praksi održivog razvoja u mikro i malim poduzećima korištenjem *Graded Response Model-a* (GRM) s Bayesovim pristupom Teoriji odgovora na zadatke. Rezultati dobiveni na temelju uzorka od 506 sveučilišnih studenata u Peruu upućuju na to da je razvijen važeći mjerni instrument. Rad završava iznošenjem metodoloških i menadžerskih doprinosa.

ment tools for the economic, social and environmental management of micro and small enterprise (MSE). In this context, this study aims to validate an instrument to assess perceptions of sustainable development practices by MSEs by means of a Graded Response Model (GRM) with a Bayesian approach to Item Response Theory (IRT). The results based on a sample of 506 university students in Peru, suggest that a valid measurement instrument was achieved. At the end of the paper, methodological and managerial contributions are presented.

1. INTRODUCTION

Corporate social responsibility (CSR) is an issue of growing concern in the political, business and educational fields (Apospori, Zografos & Magrizos, 2012). In most economies around the world, micro and small enterprises (MSE) constitute the majority of enterprises, in some cases more than 90 percent. They significantly contribute to job creation and income generation while also meeting the needs in certain marketplaces which are not very attractive to big corporations (Andriani, Biasca & Rodriguez, 2004; Apospori et al., 2012; Dahl, 2011; Holt, 2011; Morsing & Perrini, 2009). According to Apospori et al. (2012), MSEs play a fundamental role in big companies' supply chains and their impact on social and environmental practices is significant. Furthermore, the authors add that it is important to study CSR in light of MSEs' particular features, and not in the same context in which big corporations are analyzed.

Researchers list the integration of environmental and social issues into the company's core business processes as one of the main challenges (Schaltegger & Wagner, 2011; Jucan & Jucan, 2010). Businesses which contribute the most to sustainable development, in both society and the economy, are those seeking solutions to environmental and social problems, something that requires becoming innovative in terms of sustainability. The size of the company does not exempt it from being socially responsible or from contributing to improve the environment (Schaltegger & Wagner, 2011).

In this sense, Enderle and Tavis (1998) state that some of the challenges faced by society faces include social, economic and environmental dimensions that are mingled together and affect companies. Therefore, a balance among social, economic and ecological responsibilities must be found, as well as measures to capture corporate accountability along those three dimensions. Cohen, Brock and Mitchel (2008) add that defining the scope of dependent variables in corporation performance research, such as the concern for successfully reaching economic, social and environmental goals, is to go beyond financial result measurements when entrepreneurship is being studied.

However, the indices used to measure the outcomes of CSR actions have been developed mainly for big corporations (Apospori et al., 2012), for example, the indices elaborated by the Global Reporting Initiative (2008) and the ONG Perú 2021 (2008), are not adapted to and do not reflect the needs of MSEs; therefore, they cannot be applied to that context. Consequently, the relevance of proposing measures (measurement instruments) adapted to the reality of MSE economic, social and environmental management must be recognized. The need to assess MSEs' effective and sustainable performance has scarcely been met and few studies that contribute to their definition on an exploratory level have been found in the literature (Hernani & Hamann, 2013) or in in-depth interviews (Silva & Chauvel, 2011). Therefore, we still do not have studies at a confirmatory level that can help us validate the psychometric characteristics of the instruments measuring the perception of the sustainable development of MSEs.

In response to that lack of information, this study aims to validate an instrument to measure perceptions of MSE sustainable development practices using a Graded Response Model (GRM) of Item Response Theory (IRT) Bayesian approach, based on the scale developed by Hernani and Hamann's (2013). As opposed to Classical Testing Theory (CTT), IRT allows for assessment of the instrument's psychometric characteristics by focusing on item properties rather than on the properties of the test as a whole, making instrument creation more robust (Muñiz, 1997). The use of IRT, and particularly of GRM, allows us to analyze the items using an ordinal scale, which enables us to overcome the limitations of previous studies that used the classical approach which assumes interval properties for the items, which do not match the nature of the questionnaire. Despite these advantages, there

are still few studies that use this approach in the field of administration, as can be seen in Bazán et al. (2011). On the other hand, the Bayesian approach, which differs from the classical one mainly by considering parameters as random variables characterized by an *a priori* distribution (Ntzoufras, 2009), enables better estimation because of the possibility of obtaining model parameter distributions *a posteriori* and calculating confidence intervals.

2. THEORETICAL BACKGROUND

This section presents the three pillar concepts for the investigation: sustainable development, focusing on its origins, definition and scope; micro and small enterprise (MSE); and the CSR framework. We also present studies that define and justify sustainable development dimensions before proposing a perception measurement instrument for the MSE context.

2.1. Sustainable development

The concept of sustainable development appeared for the first time at a global level in 1980, in the context of the Global Strategy for Conservation initiative, in a document developed by the conservationist organization Global Union for Nature. The document stated that "sustainable development demands that social and ecological, besides economic, factors are taken into account, regarding living and nonliving resources; and that the advantages and disadvantages of alternative actions in the short and long term are also considered" (Galarza, Gómez & Gonzáles, 2002, p. 11). Later, a bulletin by the Global Commission on Environment and Development called Our Common Future defined Sustainable Development as: "The one that satisfies present needs without compromising future generations' ability to meet their own needs" (Galarza et al., 2002, p. 11).

From those definitions, the dimensions of sustainable development can be described as: economic, social and environmental. The first dimension comprises economic development, which enables companies to build structures that can foster progress in the communities, regions and countries where they operate (Conde, 2003). This dimension seeks to create value for shareholders or owners of the companies since it will generate new jobs and competitiveness between companies. On the other hand, the creation of value added products will also be necessary due to the investment in new technologies which will, at the same time reduce costs, generate new investments and improve the quality of jobs (Galarza et al,. 2002).

The social dimension addresses social welfare at all its levels, from workers to local communities and society at large, including enterprises (Centro de Investigación en Geografía Aplicada e Instituto de Estudios Ambientales; Pontificia Universidad Católica del Perú, 2008). This dimension also considers job stability, the protection of the fundamental rights of employees and the improvement of quality of life (Conde, 2003). The environmental dimension seeks the preservation of biodiversity and ecosystems, the decrease in the use of non-renewable resources, the adequate use of renewable and sustainable resources over time, as well as the control of emissions and residues (Conde, 2003). Another aspect included in this dimension is the minimization of negative environmental impacts, since reducing the environmental impact of products throughout their life cycle should be the starting point for raising environmental awareness among people and companies, making the latter responsible for any action that causes a negative environmental impact, assuming that an ecological cost could result (Artaraz, 2002).

Accordingly, Jucan and Jucan (2010) and Schaltegger and Wagner (2011) recently added that sustainable development requires a successful integration of environmental, social and economic goals, both for present and future generations. Also, integrating environmental and social

issues into the core processes of the enterprise is considered one of the main corporate goals. Besides, the companies which contribute most to economic and societal sustainable development are those that seek solutions to social and ecological problems, something that demands becoming innovative in terms of sustainability. For Albareda and Ibáñez (2010), CSR and sustainable development are areas of the complex function of globalization.

To face future challenges, concerted action is needed from government and society, along with a reframing of the interactions between business practices, political systems and society in general (Juncan, 2010). This would require that governments take responsibility to ensure that present rules consider each one of the three pillars of sustainable development: economic, social and environmental. This will lead governments to collaborate with the private sector, communities, consumers and unions, among others; to develop and apply best practices to succeed in areas such as health, environment, and equal opportunities; and to promote open and constructive dialog.

2.2. Micro and small enterprises and corporate social responsibility

In most countries, MSEs constitute the majority of enterprises and, in some cases account for more than 90 percent of the market. They contribute significantly to job creation, revenue generation and fulfill needs in some marketplaces that are not very attractive to big corporations (Andriani et al., 2004; Apospori et al., 2012; Dahl, 2011; Holt, 2011; Morsing & Perrini, 2009). In many cases, they give birth to future important entrepreneurs and are an important channel of training and development for millions of people. A natural conclusion is that their competitiveness is a strategic element for the sustainable development of any country.

According to Jenkins (2009), the size of a company does not exempt it from being socially responsible or from contributing to improvement of the environment. Companies are accountable for their actions and for the information they convey not only to their owners, but also to workers, clients, suppliers, government and society (Corral, Isusi & Vives, 2005). Unions, government agencies and universities should promote the ethical development of social and ecological responsibility practices among MSEs within their capabilities, so that they can contribute to sustainable development. Apospori et al. (2012) stress that CSR practices can be part of a consistent strategic business model and the result of long term commitment.

Dahl (2011) points out that embracing CSR depends on the size of enterprises, as the costs and returns involved in such activities may be different, causing disproportional management costs that render them more costly for smaller companies. For example, writing CSR reports demands almost the same amount of time in big enterprises as it does in small ones, but in small enterprises it requires a bigger share of the available personnel. Besides, small enterprises normally operate on stricter margins than big ones, and are more involved in day-to-day activities than in investing in CSR activities which may not pay off in the long term (Apospori et al., 2012; Dahl, 2011; Fenwick, 2002). Furthermore, small business managers tend to be in charge of a large variety of tasks and, thus, have less available time than their counterparts in large corporations to design a long-term strategy (Dahl, 2011). Lepoutre and Heene (2006) recognize MSE limitations (time, resources) to commit to CSR, but point out that they may overcome those limitations.

Corporate Social Responsibility efforts help improve company reputation, and this is more interesting to large corporations, which tend to have more visibility in capital markets, than it is to small ones that do not participate in stock markets and are, thus, less exposed to CSR evaluations (Dahl, 2011). This is even more applicable in the case of environmental issues, as a number of small enter-

prises are not aware of the impact they have on the environment. Furthermore, small businesses are less closely examined by interest groups and the media than large corporations, resulting in less pressure to behave responsibly (Dahl, 2011). According to Albinger and Freeman (2000) and Graafland and Smid (2004), companies that exhibit socially responsible behavior are more appealing to job candidates. Besides, CSR practices may lead to high confidence and cooperation levels in the workplace, collective problem solving, higher job satisfaction and lower employee turnover.

It is evident that CSR is a key condition in the maintenance of a global market economy; governments, corporations and civil society will have to accept and apply such practices if they want to meet 21st century demands. It is precisely why it is important to maximize the role played by CSR in MSEs, create measurement tools adapted to their reality to help their economic, social and environmental management, and assess their effective and sustainable performance.

2.3. Origins of sustainable development variables in MSE

Murillo and Lozano's (2006) study focusing on French and Spanish small businesses showed that most of them justify CSR efforts for activities that have an effect on their market share. However, the authors noted that none of the companies could demonstrate the impact of their CSR efforts through a quantitative measure that showed their effective results on the balance sheet. In those small businesses, the most common assessment of CSR benefits was a subjective evaluation. In this sense, a study by Besser and Miller (2001) found that, among the managers of 675 small businesses in Iowa (USA), a strong belief in commitment to CSR was a subjective measure of their companies' success.

In their pioneer study on the conceptual framework for enterprises, Enderle and Tavis (1998) considered the notions of social responsibility and the balance among economic, social and environmental issues. Specifically, they proposed various possibilities to assume corporate responsibilities and compared them in three enterprises (firm A, firm B, firm C). The authors concluded that all three firms struck a balance between economic, social and environmental responsibilities, but in different ways, depending on both their own mission and their business environmental policy. This implies a circular interrelationship among these issues.

Enderle and Tavis (1998) mentioned that corporate responsibility depends on how the corporation's role and purpose are conceived in relation to society. But if society is divided into distinct and separate domains such as the economic, the political and the sociocultural, the role of the corporation can be seen in purely economic terms, and its responsibility limited to its financial purpose. In reality, the various domains, albeit autonomous to a certain point, are interconnected, so the company concept should reflect a more comprehensive understanding of society (Enderle & Tavis, 1998). This means using measures that capture the concept of business responsibility along the three dimensions, and implies that CSR goals and measures should be defined in the planning, execution and control sequence through a three-stage process: strategic positioning; resource commitment; and evaluation. Decisions and activities in each phase should be planned focusing the long range.

Society's numerous challenges include the above mentioned dimensions. They cannot be separated from one another, and all of them affect companies, which should not act as purely economic organizations. As moral, responsible players, they have to deal with those challenges on the corporate level, balancing social, economic and environmental responsibilities (Enderle & Tavis, 1998).

On the other hand, Enderle and Tavis (1998) stress that the CSR measures defined by companies should be able to promote corporate responsibilities in the economic, social and en-

vironmental dimensions, as well as offer a performance evaluation tool for each one of these. In the economic dimension, assessing income is paramount, for it adequately reflects productivity and the long term impact of the commitment of resources. In the environmental dimension, there is concern among nations to curb the excessive consumption of natural resources as inputs, and about the impact of waste as an output; the responsibility for monitoring consumption and reducing residues lies with each multinational company. In the social dimension, measurement is directed towards respecting diversity, interacting with local players and societies, and respecting cultures.

In their research of the business literature to identify CSR measures for small businesses, Cohen *et al.* (2008) found that a key idea among entrepreneurs is the search for profitability, which leads to the exclusion of other factors that impact the firm. The authors propose a typology for business value creation reflecting the concept of the triple baseline (dimensions): economic, social and environmental, following Enderle and Tavis's view. Entrepreneurs' economic, social and environmental motivations not only define seven types of corporate motivations and objectives, but also seven sets or domains of business value creation (or efficiency metrics) that can be used to evaluate whether those objectives have been met.

Establishing an instrument that reflects this triple account and its results (dimensions) is an important and necessary step for both scholars and practitioners. It may provide a normative guide to go beyond measuring financial results when studying the phenomenon of entrepreneurship, by introducing a viable, multidimensional solution to help create effective measures to assess company performance considering social and environmental criteria (Cohen et al., 2008).

The conceptual framework proposed by Enderle and Tavis (1998) introduces the notion of balance in economic, social and environmental responsibility, and implies a circular interrelation among those dimensions, in which one dimension cannot be overemphasized to the detriment of the others. This balance can help to reduce the gap between internal and external company evaluation, assuring that a long term view is firmly integrated into all planning efforts. This means that, whatever model is used, the balance between today's and tomorrow's decisions should be explicit and reflected in the measurements.

Studies by Cohen et al. (2008) and by Enderle and Tavis (1998) also mention the three dimensions as areas of corporate responsibility. Regarding the economic dimension, they propose items related to productivity and earnings distribution such as: benefits maximization, increased productivity, maintenance/increase of owner and investor wealth, respect for suppliers, fair competition, job creation and retention, fair wages, provision of social benefits, employee training, and serving customers. The social dimension comprises respect for laws, social habits and cultural inheritance; and a selective participation in cultural and political life. Finally, the environmental dimension refers to commitment to sustainable growth, consumption of fewer natural resources, and disposal of less waste into the environment.

Apospori et al. (2012) point out that measuring the results of CSR practices is a difficult issue, as most indices are not yet being measured on a constant basis and do not reflect social impacts precisely. The authors note that many indices designed to measure businesses' social responsibility, such as FTSE4Good, have been developed mainly for big corporations and cannot be applied in the context of small businesses. In addition, those indicators are not adapted to, nor do they reflect the needs and demands of small businesses; consequently, their application to the assessment of CSR actions performed by small companies is questionable.

As shown by the studies presented here, there is a need to develop CSR performance measures (measurement instruments) adapted to the reality of small businesses in terms of their economic, social and environmental management. The measures proposed so far in the literature are still

incipient, even at an exploratory level (Hernani & Hamann, 2013; Silva & Chauvel, 2011). A new measurement tool could profit from a solid theoretical-practical basis if it was developed inside universities (Dale & Newman, 2005). Therefore, in this study we intended to validate an instrument at a confirmatory level, and describe its psychometric properties in a more robust way, so that the perceptions of MSE sustainable development practices may be measured based on the dimensions proposed by Enderle and Tavis (1998) in their pioneer work.

3. METHODOLOGY

This is an exploratory-descriptive study (Hernández, Fernández-Collado & Baptista, 2006), the objective of which is to refine and validate an instrument to measure MSE sustainable development actions.

3.1. Sample

The sample consists of 506 business undergraduate students from five universities in the city of Lima (Peru): 365 students from private universities and 141 from a public university. In total, 50.8 percent of the respondents were male, but in the private university subgroup they accounted for 56.4 percent. Conversely, only 36.3 percent of the students from the public university were male. While 49.2 percent of the total respondents were female, among students from the public university women were the majority (63.7 percent) versus 43.2 percent in the private ones. For further detail, see Hernani and Hamann (2013), who designed and validated a scale at an exploratory level to measure the perceptions of students in relation to the sustainable development of the MSE.

The sample, according to Malhotra (2012), is not probabilistic as the discoveries are not generalizable to the population. In addition, it is a convenience sample because it relies on the judgment of the investigator to obtain the sample elements. In this way, the application of this instrument to university students is justified, since it is a homogenous sample, as is necessary for a scientific study, and also because they have the knowledge and management skills to implement it in an MSE context.

3.2. Instrument

The instrument used in the study was developed by Hernani and Hamann (2013) to measure future graduate professionals' perception of sustainable development efforts by MSEs. The instrument contains 46 items related to the three constructs (environmental, social and economic) of MSE sustainable development (see Annex 1). Each item was measured using a 5-point Likert scale, ranging from 1 (complete disagreement) to 5 (complete agreement). The instrument was previously validated through an analysis based on classical testing theory (Muñiz, 1996), and Cronbach's Alpha coefficients were 0.90 for the environmental dimension, 0.92 for the social and 0.97 for the economic dimension.

3.3. Procedure

The psychometric analysis of the scale developed by Hernani and Hamann (2013) was done considering the following: (1) evaluation of the scale dimensionality; (2) evaluation of the items according to a Bayesian graded response model.

3.3.1. Scale dimensionality analysis

A Bayesian Factor Analysis (BFA) was done to assess the occurrence of multidimensionality in the latent trait and to verify if the business undergraduate students' perceptions of MSE sustainable development practices were represented

along the economic, social and environmental dimensions mentioned by Enderle and Tavis (1998). The factor analysis was executed through the *MCMCFactanal* function of *MCMCPack* (Martin, Quinn & Park, 2011) of the R Statistics software. Specifically, both the scores and the factor loadings were assumed a priori to have normal distribution, while for the responses an Inverse Gamma distribution was assumed *a priori*.

The *MCMCFactanal* function allows the generation of a sample of the model's distribution *a posteriori*, using the Markov–Monte Carlo (MCMC) chain simulation with Gibbs's standard sampling algorithm. BFA can be considered a confirmatory factor analysis since the results yielded by an exploratory factor analysis (EFA), or any theoretical suppositions about the instrument's item groupings one wishes to verify, can be used as restrictions for the factor loadings. For the EFA, we used the results from the pilot study by Hernani and Hamann (2013), which yielded three dimensions with accumulated variance of 42.7 percent – a result considered by Hair, Anderson, Tatham and Black (1999) as satisfactory. In order to obtain the *a posteriori* probability distributions for factor loadings and the unidimensionality, a total of 100,000 simulations were run with an initial specification of 1,000 parameter vectors. Parameter vectors were recorded for every 10 simulations. The convergence of the simulations was checked through Geweke's (1992) and Heidelberger-Welch's (1981) diagnostics in the CODA package in R software (Development Core Team, 2012). It is worth mentioning that a review of the works by Brooks & Roberts (1998) and Cowles & Carlin (1996) is highly recommended for a further discussion of convergence criteria.

3.3.2. Item evaluation

The approach used to validate the instrument was Item Response Theory (IRT) (Lord & Novick, 1968). According to Muñiz (1997), IRT aims to evaluate the instrument's psychometrics characteristics focusing on item properties, rather than on the properties of the test as a whole, as opposed to Classical Testing Theory (CTT).

Pasquali (2009) describes the main characteristics of IRT as: (1) the subject's response to an item is explained as a function of a set of factors or latent traits (capabilities, skills, etc.); and (2) the relationship between the subject's responses and the latent traits can be described through a monotonous growing equation called an Item Characteristic Curve.

The main advantages are the possibility of obtaining measurements that do not vary as a function of the measurement instrument, as it has psychometric properties that do not depend on the profile of the subjects under evaluation (Muñiz, 1997). Another advantage is that IRT permits the estimation of distinct values for capabilities in situations that would yield similar values using traditional techniques, such as factor analysis or summated scales (Sabbag, Bernardi, Goldszmidt & Zambaldi, 2010). Also, the estimates for the various model parameters allow the determination of items which should remain in the final version of the instrument. However, the focus of IRT does not invalidate or contradict that of CTT, but rather provides a larger scope due to its focus on the items. For this reason, the classical indices are also provided in this study.

Because the items were measured on a Likert-type scale, Samejima's (1969) Graded Response Model (GRM) was used, as it was specifically designed to be utilized with polytomous data using an ordinal scale.

According to Azevedo (2003), if it is assumed that the categories of a certain item j can be ranked in growing order and denoted as, being the number of categories of item number *j*, the probability that an individual *i* who has a certain level of the latent trait chooses category *k* or above of item *j* is calculated as:

$$P^{+}_{i,j,k} = P(Y_{ij} \geq k|\theta_i) = \frac{exp[a_j(\theta_i - b_{jk})]}{1 + exp[a_j(\theta_i - b_{jk})]},$$

$$con\ i = 1, \dots, n, j = 1, \dots, J\ y\ k = 0, 1, \dots, m_j,$$

Where:

- Y_{ij}: response by individual number *i* to item *j*.
- θ_i: individual *i*'s latent trait (sustainable perception).
- a_j: discrimination parameter (scale) for item j.
- b_{jk}: difficulty parameters for response alternative *k* for item *j*, also known as transition or threshold parameters.

The category ranking definition must be:

$$b_{j1} \leq ... \leq b_j m_j$$

meaning that the difficulty levels of a certain item's categories are ranked.

From that, the probability that an individual *i* responds to category *k* for item i is calculated as:

$$P_{i,j,k} = P_{j,k}(\theta_i) = P^{+}_{j,k}(\theta_i) - P^{+}_{j,k+1}(\theta_i) \quad =$$

$$= \frac{exp[a_j(\theta_i - b_{jk})]}{1 + exp[a_j(\theta_i - b_{jk})]} - \frac{exp[a_j(\theta_i - b_{jk})]}{1 + exp[a_j(\theta_i - b_{j(k+1)})]}$$

The likelihood for the response vector *y* is given by:

$$L(\boldsymbol{y}|\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta}) = \prod_{i=1}^{n} \prod_{j=1}^{J} \prod_{k=0}^{m_j} P_{i,j,k}^{y_{ijk}}$$

In the Bayesian approach, a standard normal distribution for each latent trait θ_i is assumed, namely $\theta_j \sim N(0,1)$.

The *a priori* distributions used for the discrimination parameters were:

$$a_j \sim N_{+}(m_a, s_a^2)$$

where m_a and m_a are constants specified before the analysis, $N + (m, s^2)$ denotes a normal distribution with mean *m* and variance s^2.

Following Curtis (2010), the difficulty parameters were defined as $\kappa_{jk} = a_j * b_{jk}$. These new parameters still follow the order restriction $\kappa_{j1} \leq ... \leq \kappa_{jm_j}$ and can be induced by defining the auxiliary parameters $\kappa^*_{j1}, ..., \leq \kappa^*_{jm_j}$, without restrictions, so that:

$$\kappa^*_{jk} \sim N(m_k, s^2_k)$$

The apriori distributions of the transition parameters for item *j* are then computed setting κ_{jk} as equal to the rank statistic *k* of the auxiliary variables $\kappa^*_{j1}, ..., \leq \kappa^*_{jm}$ for item *j*, namely:

$$\kappa_{jk} = \kappa^*_{j,[k]}$$

In order to get the *a posteriori* probability distributions of the GRM parameters in each dimension, a total of 10,000 simulations were performed with an initial specification of 4,000 parameter vectors. The parameter vectors in every 10 simulations were recorded, considering three chains, using the bugs function in R2WinBUGS in the R *Statistics* software package. Curtis' (2010) GRM syntax, performed using WinBUGS (Lunn, Thomas, Best & Spiegelhalter, 2000), was modified to yield the difficulty parameters (*b*). As for the convergence of the simulations, besides running the diagnostic tools in BFA, the effective sample size was evaluated. Another test run was Gelman and Rubin's (1992) diagnostic, which is displayed as *Rhat*. *Rhat* values close to 1 indicate good convergence.

4. RESULTS

First the scale dimensionality was assessed on a confirmatory basis using BFA with MCMC, considering the three factors found in the EFA reported by Hernani and Hamann (2013), along with Enderle and Tavis's (1998) definition. Previously, each item was restricted to the factor it should be associated with. The factor loadings yielded by the *MCMCFactanal* software are shown in Table 1. The results confirm the presence of the three dimensions previously considered by the cited authors, since most items are strongly associated to each dimension, except items 2 and 43, with factor loadings of 0.36 and 0.35 respectively. Regarding MCMC convergence tests, the *p* values for the Z distribution in Geweke's

Table 1: Results of factor analysis and item analysis

Dimension		Classical indices				GRM indices d				
		M	S	CITC	FL	b1	b2	b3	b4	a
Environmental (Alpha = 0.90)	1	2.51	0.88	0.43	0.43	-2.27	0.04	2.70	5.16	0.94
	2	2.56	0.88	0.36	0.36	-2.76	0.00	2.82	5.64	0.82
	3	2.19	0.87	0.58	0.59	-1.06	0.59	2.32	3.83	1.61
	4	2.25	0.91	0.62	0.64	-1.11	0.54	1.85	3.65	1.73
	5	2.34	0.94	0.63	0.66	-1.21	0.38	1.66	3.35	1.74
	6	2.71	0.91	0.49	0.51	-2.21	-0.39	1.56	4.51	1.18
	7	2.58	0.97	0.67	0.69	-1.42	-0.10	1.30	2.88	1.90
	8	2.24	0.92	0.59	0.63	-0.98	0.43	2.15	3.82	1.61
	9	2.66	1.07	0.49	0.53	-1.69	-0.07	1.42	2.84	1.27
	10	2.14	0.95	0.65	0.73	-0.70	0.52	1.80	3.00	2.13
	11	2.20	0.87	0.63	0.7	-0.97	0.50	1.99	3.64	2.02
	12	2.17	0.92	0.71	0.79	-0.77	0.50	1.62	2.79	2.69
	13	2.07	0.88	0.67	0.75	-0.63	0.59	1.95	3.29	2.41
	14	2.32	0.92	0.64	0.72	-1.03	0.25	1.76	3.12	2.10
Social (Alpha = 0.92)	15	2.50	1.00	0.57	0.6	-1.48	-0.02	1.42	3.56	1.46
	16	2.48	0.92	0.63	0.66	-1.47	0.01	1.48	3.40	1.77
	17	2.85	1.02	0.57	0.6	-2.06	-0.47	0.74	2.97	1.48
	18	2.78	0.97	0.64	0.67	-1.88	-0.38	0.94	2.61	1.78
	19	2.42	0.92	0.63	0.66	-1.38	0.11	1.57	3.31	1.75
	20	2.50	0.95	0.6	0.63	-1.54	0.04	1.42	3.32	1.63
	21	3.08	1.00	0.61	0.64	-2.35	-0.80	0.50	2.22	1.65
	22	2.46	0.91	0.6	0.64	-1.56	0.04	1.53	3.62	1.68
	23	2.75	1.04	0.65	0.69	-1.59	-0.36	0.96	2.21	1.88
	24	2.67	0.98	0.67	0.72	-1.59	-0.19	1.08	2.44	2.05
	25	2.70	0.94	0.65	0.7	-1.72	-0.28	1.08	2.61	1.99
	26	2.58	0.91	0.64	0.68	-1.69	-0.07	1.25	3.26	1.90
	27	2.42	0.88	0.59	0.63	-1.56	0.14	1.81	3.83	1.59
	28	2.33	1.06	0.64	0.69	-0.93	0.28	1.38	2.54	1.88
	29	2.48	0.98	0.56	0.6	-1.51	0.08	1.53	3.21	1.48
	30	2.66	1.03	0.59	0.63	-1.59	-0.24	1.07	2.84	1.60
	31	2.66	0.95	0.69	0.73	-1.60	-0.23	1.04	2.70	2.13
Economic (Alpha = 0.87)	32	2.50	1.00	0.50	0.54	-4.16	-2.52	-0.79	1.07	1.21
	33	2.48	0.92	0.54	0.6	-3.59	-1.58	0.34	2.43	1.47
	34	2.85	1.02	0.49	0.52	-3.11	-1.34	0.09	1.88	1.25
	35	2.78	0.97	0.39	0.41	-2.56	-0.30	2.05	4.54	0.85
	36	2.42	0.92	0.60	0.68	-1.75	-0.43	0.74	1.96	1.83
	37	2.50	0.95	0.60	0.65	-3.48	-2.09	-0.63	0.93	1.65
	38	3.08	1.00	0.62	0.71	-2.38	-1.04	0.09	1.23	2.01
	39	2.46	0.91	0.45	0.46	-4.05	-2.25	-0.51	1.56	0.99
	40	2.75	1.04	0.54	0.59	-2.52	-0.72	1.29	3.02	1.48
	41	2.67	0.98	0.47	0.5	-4.39	-2.99	-1.29	0.71	1.07
	42	2.70	0.94	0.50	0.56	-3.12	-1.45	0.13	2.03	1.27
	43	2.58	0.91	0.33	0.35	-4.11	-1.59	1.16	4.27	0.72
	44	2.42	0.88	0.53	0.62	-1.65	-0.11	1.66	3.11	1.63
	45	2.33	1.06	0.60	0.67	-2.12	-0.49	1.25	2.93	1.87
	46	2.48	0.98	0.54	0.57	-2.80	-1.33	0.22	1.94	1.31

Note: Alpha= Cronbach's Alpha; M=mean; S=standard deviation; CITC=corrected item-to-total correlation; FL=factor Loading; bi=difficulty parameters; a=discrimination parameter.
Source: developed by the authors based on the data.

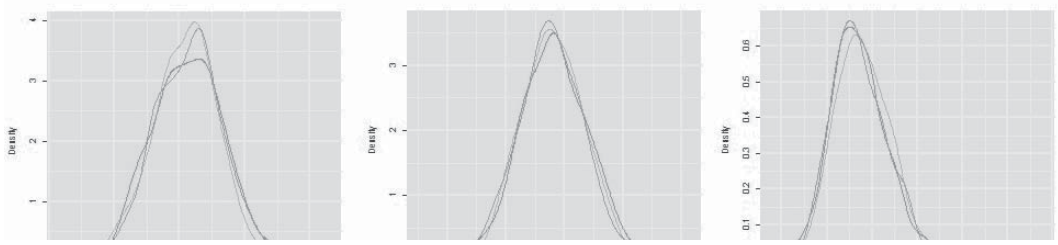
convergence diagnostic were higher than 0.05, showing no evidence against convergence. Heidelberger and Welch's (1981) stationary distribution and halfwidth tests indicated that all factor loadings stood both tests.

The items were analyzed using GRM with a Bayesian approach in order to evaluate the instrument's psychometric properties together with CTT. Regarding classical CTT indicators, the analysis included the means (Me), standard deviations (S), the corrected item-to-total correlations (CITC), factor loadings and Cronbach's Alpha coefficients (Alpha). In the case of GRM, the *a posteriori* distributions of discrimination parameters *a* and difficulty parameters *b* were computed. Table 1 shows both the classical indices and the *a posteriori* distribution means for the GRM parameters.

As for the convergence of MCMC results, it is necessary to point out that all the tests run to evaluate the BFA were also applied to GRM, showing adequate results. Also, the *Rhat* values for all the estimated parameters were assessed, showing values close to 1 in all cases. Although Table 1 only shows the mean values for the parameters, the advantage of choosing the Bayesian approach is its ability to estimate *a posteriori* the model parameters distribution (as can be seen in Figure 1). This renders the analysis more robust as it can estimate, for example, confidence intervals. After running the convergence tests, the next step was to identify which items should remain in the final version of the scale, following Thissen's (1986) criterion of considering statements with discrimination scores below 0.50 as candidates for elimination. As no item violated this guideline, the decision was made to keep all the items in the questionnaire.

Besides validating the instrument, the GMR estimated parameters can be used to characterize the different dimensions of perception of MSE sustainable development practices. For any particular item, parameter *a* indicates how much discriminating power it has, while parameter *b* signals on which level of the perception scale the item has more discriminating power. In the environmental dimension, items 12 ("MSEs offer their employees training on environmental issues to reinforce their ecological awareness") and 13 ("MSEs develop environmental education campaigns for employees' families and for the community nearby") are the ones with the highest discriminating power. On the other hand, although items 1 ("MSEs save energy to preserve the environment") and 2 ("MSEs do not waste water in their manufacturing processes") have lower discriminating power in general, they are the ones that best help in discriminating individuals with a very negative or very positive environmental perception. All the remaining parameters can be used in the same way to characterize the other dimensions.

Figure 1: A posteriori distribution of the discrimination parameter for item 1 and difficulty parameters 2 and 4



Source: developed by the authors based on the data using the *mcmcplot* function.

5. DISCUSSION AND FINAL RESULTS

This study aimed to validate the exploratory instrument developed by Hernani and Hamann (2013) to measure the perceptions of MSE sustainable development efforts. Based on the results discussed here, we can conclude that the instrument was validated with the application of an IRT polytomous model (Graded Response Model) based on Bayesian inference, as a complement to CTT.

As for the psychometric validation of the instrument, this paper offers a methodological contribution since IRT is seldom utilized in business studies. Furthermore, this methodological proposition used the Bayesian approach because of its advantages in comparison to classical estimation testing. Since parameters are considered as random variables in IRT, it allows the estimation of confidence intervals and *a posteriori* distributions that include previous knowledge, helping to create instruments with higher validity and precision. It is equally important to stress that two free software packages, R and WinBUGS, were used in this study. They offer the advantage of allowing more freedom to manipulate their codes according to the study's characteristics and, thus, improve the model estimation. Therefore, they are better suited for investigation purposes than are commercial software packages, which are usually more restricted.

From a managerial point of view, this analysis offers instruments that, considering respondents' different characteristics (skill levels), are able to identify real perceptions, thereby helping to define adequate policies to manage MSE sustainable development along its three dimensions.

Offering a valid and reliable instrument to measure the parameters for each item, as shown here, is a great contribution of academia to the corporate world. As an example, relating to difficulty parameter *b*, it was shown that the economic dimension items "MSEs produce a positive impact in the country's economy" and "Supporting the development of MSEs creates value for the country's economy" both have a positive perception in terms of CSR practices. On the other hand, the items "MSEs save energy (electricity, fuel) to preserve the environment" and "MSEs do not waste water in their manufacturing processes" both had a very negative perception. This suggests that the use of natural resources has been neglected by MSE, and this suggestion could lead to the promotion of preservation policies by the government.

For future investigation, the multidimensional focus of Item Response Theory (IRT) under the Bayesian perspective could be further explored.

REFERENCES

1. Albareda, L., & Ibañez, J. (2010). Corporate contributions to global governance through corporate social responsibility: the UN Global Compact and the Global Reporting Initiative. *7th Pan-European International Relations Conference Standing Group on International Relations Stockholm*. Stockholm.
2. Albinger, H., & Freeman, S. (2000). Corporate social performance and attractiveness as an employer to different job seeking populations. *Journal of Business Ethics*, *28*(3), 243-253.
3. Andriani, C., Biasca, R., & Rodriguez, M. (2004). *Un nuevo sistema de gestión para lograr PYMES de clase mundial*. 2 ed. México: Norma.
4. Apospori, E., Zografos, K., & Magrizos, S. (2012). SME corporate social responsibility and competitiveness: a literature review. *International Journal of Technology Management*, *58*(1/2), 10-31.

5. Artaraz, M. (2002). Teoría de las tres dimensiones del desarrollo sostenible. *Ecosistemas, 2*(11).
6. Azevedo, C. (2003). *Método de estimação na teoria de resposta ao item* [Documento de trabajo]. São Paulo, SP: Instituto de Matemática e Estatistica da Universidade de São Paulo.
7. Bazán, J. L., Mazzon, J., & Hernani, M. (2011). Clasificação de modelos de resposta ao item policotômicos com aplicação ao marketing, *Revista Brasileira de Estatística, 72*(235), 7-39.
8. Besser, T., & Miller, N. (2001). Is the good corporation dead? The community social responsibility of small business operators. *Journal of Socio- Economics*, *30*(3), 221-241.
9. Brooks, S., & Roberts, G. (1998). Assessing convergence of Markov chain Monte Carlo algorithms. *Statistics and Computing, 8*, 319-335.
10. Centro de Investigación en Geografía Aplicada e Instituto de Estudios Ambientales; Pontificia Universidad católica del Perú. (2008). La tierra vista desde el cielo. Recuperado el 04 de febrero, 2008, de www.pucp.edu.pe/idea
11. Cohen, B., Smith, B., & Mitchel, R. (2008). Toward a Sustainable Conceptualization of Dependent Variables in Entrepreneurship Research. *Business Strategy and the Environment Bus. Strat. Env., 17*, 107-119.
12. Conde, J. (2003). *Empresa y medio ambiente*: *hacia la gestión sostenible*. Madrid: Nivola libro y ediciones.
13. Corral, A., Isusi, I., & Vives, A. (2005). *Responsabilidad Social de la empresa en las PYMES de Latinoamérica*. España: IKEI.
14. Cowles, M. K., & Carlin, B. P. (1996). Markov chain Monte Carlo convergence diagnostics: a comparative review. *Journal of the American Statistical Association, 91*, 883-904.
15. Curtis, M. (2010). Bugs Code for Item Response Theory. *Journal of Statisticial Software, 36*(1), 1-34.
16. Dahl, R. (2011). Little big firms? Corporate social responsibility in small businesses that do not compete against big ones. *Journal of Business Ethics*, *20*(3), 30-44.
17. Dale, A., & Newman, L. (2005). Sustainable development, education and literacy. *International Journal of Sustainability in Higher Education*, *6*(4), 351-362.
18. Development Core Team (2012). *R: A language and environment for statistical computing*. Recuperado el 01 marzo, 2012, de http://www.R-project.org/
19. Enderle, G., & Tavis, L. (1998). A Balanced Concept of the Firm and the Measurement of Its Long-term Planning and Performance. *Journal of Business Ethics*, *17*(11), 1129-1144.
20. Fenwick, T. (2002). Lady, Inc.: Women learning, negotiating subjectivity in entrepreneurial discourses. *International Journal of Lifelong Education*, *21*(2), 162-177.
21. Galarza, E., Gómez, R., & Gonzáles, L. (2002). *Ruta hacia el desarrollo sostenible del Perú* [Documento de trabajo]. Lima: Centro de Investigación de la Universidad del Pacífico (CIUP).
22. Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science, 7*(4), 457-511.
23. Geweke, J. (1992). Evaluating the accuracy of sampling–based approaches to the calculation of posterior moments. In: J. O. Berger, J. M. Bernardo, A. P. Dawid & A. F. M. Smith (eds.). *Bayesian Statistics. 4* (pp. 169-193). Oxford: Clarendon Press.
24. Global Reporting Initiative. (2008). *Guía para la elaboración de memorias de sostenibilidad*. Retrieved 28 Augost, 2008, from www.globalreporting.org
25. Graafland, J. J., & Smid, H. (2004). Reputation, corporate social responsibility and market regulation. *Tijdschrift voor Economie en Management*, *49*(2), 271-308.
26. Hair, Jr. J. E., Anderson, R. E., Tatham, R. L., & Black, W. C. (1999). *Análisis multivariante. 5*ed, Madrid: Prentice may Iberia.
27. Heidelberger, P., & Welch, P. D. (1981). A spectral method for confidence interval generation and run length control in simulations. *Comm. ACM*, *24*, 233-245.

28. Hernández, R., Fernández, C., & Baptista, P. (2006). Metodología de la Investigación. 4ed, México: McGraw-Hill.
29. Hernani, M., & Hamann, A. (2013). Percepción sobre el desarrollo sostenible de las MYPE en el Perú. *RAE - Revista de Administração de Empresas, 53*(3), 290-302.
30. Holt, D. (2011). Where Are They Now? Tracking the Longitudinal Evolution of Environmental Businesses from the 1990s. *Business Strategy and Environmental*, *20*, 238-250.
31. Jenkins, H. (2009). A business opportunity' model of corporate social responsibility for small- and medium-sized enterprises. *Journal of Business Ethics*, *18*(1), 21-36.
32. Jucan, C. (2010). Social responsibility, crisis and sustainable development. *Business Review*, *17*(1).
33. Jucan, C. N., & Jucan, M. S. (2010). Social Responsibility in Tourism and Sustainable Development. *WSEas Transactions on Environment and Development*, *10*(8), 677-686.
34. Lepoutre, J., & Heene, A. (2006). Investigating the impact of firm size on small business social responsibility: a critical review. *Journal of Business Ethics*, *67*(3), 257-273.
35. Lord, F., & Novick, M. (1968). *Statistical theories of mental test scores*. U.S.: Addison-Wesley Educational Publishers Inc.
36. Lunn, D. J., Thomas, A., Best, N., & Spiegelhalter, D. (2000). Winbugs – a Bayesian modeling framework: concepts, structure, and extensibility. *Statistics and Computing, 10*, 325-337.
37. Malhotra, N. (2012). *Pesquisa de marketing*: uma orientação aplicada. 6ed, Porto Alegre: Bookman.
38. Martin, A. D., Quinn, K. M., & Park J. H. (2011). MCMCpack: Markov Chain Monte Carlo in R. *Journal of Statistical Software, 42*(9), 1-21.
39. Morsing, M., & Perrini, F. (2009). CSR in SMEs: do SMEs matter for the CSR agenda?. *Journal of Business Ethics*, *18*(1), 1-6.
40. Muñiz, J. (1996). *Teoría clásica de los tests*. 2ed, España: Ediciones Pirámide.
41. Muñiz, J. (1997). *Introducción a la Teoria de Respuesta a los Items*. España: Ediciones Pirámide.
42. Murillo, D., & Lozano, J. M. (2006). SMEs and CSR: an approach to CSR in their own words. *Journal of Business Ethics*, *8*(3), 227-240.
43. Ntzoufras, I. (2009). *Bayesian Modeling Using WinBUGS*. New Jersey, NJ: John Wiley & Sons.
44. ONG Perú 2021 (2008). *Indicadores de RSE Perú 2021*. Retrieved 30 June, 2008, from http://www.peru2021.org/
45. Pasquali, L. (2009). *Psicometría: teoria dos testes na psicologia e na educacao*. Brasil: Editora Vozes Ltda.
46. Sabbag, P. Y., Bernardi, P., Goldszmidt, R., & Zambaldi, F. (2010). *Validação de escala para mensurar resiliência por meio da Teoria de Resposta ao Item (TRI)*. Anais do XXXIV Encontro da Associação Nacional de Pós–Graduação e Pesquisa em Administração – EnANPAD, Rio de Janeiro, RJ, Brasil, p. 15.
47. Samejima, F. (1969). Estimation of latente ability using a response pattern of graded scores. *Psychometrica Monograph, 17*, 1-100.
48. Schaltegger, S., & Wagner, M. (2001). Sustainable Entrepreneurship and Sustainability Innovation> Categories and Interactions. *Business Strategy and the Environment*, *20*, 222-237.
49. Silva, R. C. M., & Chauvel, M. A. (2011). Responsabilidade social no ensino em administração: um estudo exploratório sobre a visão dos estudantes de graduação. *Revista de Administração Pública*, *45*, 1539-1563.
50. Thiessen, D. (1986). *Multilog: Item analysis and scoring with multiple category response models* (versión 6). [Software]. Mooresville.

Appendix 1: Research instrument

No	Items
1	MSEs save energy (electricity, fuel) to preserve the environment.
2	MSEs do not waste water in their manufacturing processes.
3	MSEs take actions to reduce greenhouse gas emissions, global warming, etc.
4	MSEs deal responsibly with solid waste.
5	MSEs separate solid waste for recycling.
6	MSEs use recycled materials in their product manufacture.
7	MSEs incentive recycling internally.
8	MSEs utilize clean technologies (natural gas, solar panels).
9	MSEs know the environmental damage caused by their production activities.
10	MSEs get together with the government and communities to discuss environmental issues.
11	MSEs control their activities that cause excessive noise.
12	MSEs offer their employees training on environmental issues to reinforce their ecological awareness.
13	MSEs develop environmental education campaigns for employees' families and for the community nearby.
14	MSE support or participate in educational projects in association with environmental protection organizations.
15	MSEs offer comprehensive social benefits to their employees.
16	MSEs establish collective agreements that benefit their employees.
17	MSEs communicate their employees when they do job rotations.
18	MSEs keep health and safety programs for their employees.
19	MSEs offer their employees' families and the community preventive health programs.
20	MSEs hold professional training programs for their employees.
21	MSEs evaluate employee performance.
22	MSEs pay their employees fair wages according to their professional skills.
23	MSEs have a code of ethics.
24	MSEs give their personnel orientation and education regarding ethical principles.
25	MSEs know the concept of corporate social responsibility.
26	MSEs know and care for the needs of the community in which they are established.
27	MSEs employees join in charities or volunteer activities.
28	MSEs respect and promote intellectual property rights (copyright, patents etc.).
29	MSEs present socially-oriented proposals to government officers aiming to see them approved and implemented.
30	MSEs inform their customers about their product and service production processes.
31	MSEs follow the law and ethical norms related to marketing communications, such as advertising, publicity and sponsorship.
32	MSEs create economic value for society.
33	MSEs are considered good clients by their suppliers.
34	MSEs have access to the financial system.
35	MSEs get enough help from the government.
36	MSEs generate formal jobs.
37	MSEs produce a positive impact to the country's economy.
38	MSEs contribute to the formalization of the economy.
39	Present legislation is intended to help MSE get formal.
40	MSEs pay their lawful taxes punctually.
41	Supporting the development of MSE creates value for the country's economy.
42	MSEs are companies with high productivity in Peru.
43	MSEs get help from non-government organizations.
44	MSEs have formal and transparent accounting systems.
45	MSEs abide by the national laws in their sectors.
46	Present legislation is intended to help the creation and development of MSEs and to promote competitiveness, sustainable jobs, productivity and profitability.

Note: Environmental Dimension (items 1 to 14), Social Dimension (items 15 to 31) and Economic Dimension (items 32 to 46).