

A PSYCHOMETRIC ASSESSMENT OF THE LPME SCALE FOR THE SOUTH AFRICAN SKILLS DEVELOPMENT CONTEXT

*Maelekanyo Christopher Tshilongamulenzhe**

Abstract

A thorough examination of psychometric properties of measurement scales is necessary to ensure that these scales comply with the existing scientific conventions. This article assesses the psychometric properties of the Learning Programme Management and Evaluation (LPME) scale. A quantitative, non-experimental cross-sectional survey design was used. Data were collected from a sample of 652 respondents comprising skills development practitioners and learners/apprentices. Data were analyzed using Winsteps, SPSS and AMOS computer software. The findings show that the LPME scale meets the psychometric expectations and complies with the established scientific conventions in terms of validity, reliability, fit and unidimensionality.

Keywords: Skills Development, Exploratory Factor Analysis, Occupational Learning Programme, Management, Evaluation

**Department of Human Resource Management, University of South Africa, Pretoria, P.O.Box 392, Unisa, 0003*

Tel: +27124293724

Fax: +27866422062

1 Introduction

The South African skills shortage challenge is documented in the literature (Erasmus and Breier, 2009; SRI, 2008; Tshilongamulenzhe, 2012a; Visser and Kruss, 2009). This challenge became very glaring because of increased investment in public infrastructure over the past few years, laying bare the fact that although the funding for the infrastructure is available, there is lack of skilled people to do the construction (Sebusi, 2007). The huge skills requirements for the infrastructural developments required to facilitate the 2010 FIFA World Cup, the Gautrain Rapid Rail project, the Gauteng Freeway Improvement Programme, as well as the Eskom Electricity Supply Capacity Expansion program are some stark examples of the growth and expansion projects which placed a major drain on the state of available skills in South Africa (Townsend, 2006). The shortage of skills is prevalent across the labour market from entry-level technical occupations to management as highlighted in the 2014 National List of Occupations in High Demand (DHET, 2014).

Consequently, occupational learning programmes are publicized as a fundamental mechanism to address skills shortages in the South African context (Wildschut et al., 2012), hence vocational and occupational certification via learnership and apprenticeship programmes is at the core of the new skills creation system. An occupational learning programme is a learnership, an

apprenticeship, a skills programme or any other prescribed learning programme that includes a structured work experience component (Coetzee et al., 2012; Republic of South Africa, 2008; Van Rooyen, 2009). These programmes are inserted into a complex and increasingly bureaucratized qualifications and quality assurance infrastructure. They are administered by the Sector Education and Training Authorities (SETAs), which are in effect, a set of newly created institutions that have yet to develop capacity to drive skills development (Marock et al., 2008).

A number of challenges have been raised regarding the co-ordination and management of skills development training projects in South Africa (Du Toit, 2012), including poor quality of training and lack of mentorship. Consequently, this research seeks to address the research gap of a country-specific valid and reliable instrument to assess the effectiveness of management and evaluation practices pertaining to occupational learning programmes.

The efficacy of occupational learning programmes is reliant on the contribution of all key stakeholders from policy implementation to learner beneficiaries (Tshilongamulenzhe, 2012b). Best practice dictates that strategies relating to human resources and specifically human resource development (HRD) are enhanced when all stakeholders are able to offer their contribution and perceived opinions with regard to the efficacy of occupational learning programmes (Skinner et al.,

2004). However, Lundall (2003) maintains that occupational learning programmes are fraught with inefficiency and have a long way to go in order to prove themselves in terms of teaching and learning excellence and quality.

2 Learning programme management and evaluation challenges in South Africa

An 'Impact assessment study of the National Skills Development Strategy (NSDS) II' (Mummenthey et al., 2012) revealed the prevalence of difference in standards across the different occupational learning routes, which brought about inconsistencies regarding procedures to implement training. This was found to significantly impact on the uniformity and reliability of the outcome, resulting in confusion amongst providers and workplaces. The inconsistent implementation of workplace learning demonstrates that more guidance and improved quality assurance mechanisms are required. Further, the study (Mummenthey et al., 2012) revealed that there is a lack of structured and sufficiently monitored practical work-exposure as well as full exposure to the trade, particularly in the case of apprenticeships in the workplace. The quality checks were found to be superficial: checking policies and procedures, but not thoroughly checking what is actually happening during training. The primarily paper-based checks (sometimes adding learner interviews) were found to be insufficient and "completely missing the point" (Mummenthey et al., 2012, p. 40). A lack in subject matter expertise often reduced the process of quality assurance to a paper proof instead of actually assuring the quality of training.

However, overall alignment of theory and practice could be better achieved through setting and maintaining a consistent benchmark for training at institutional and workplace level. Minimum standards in terms of learning content and workplace exposure, together with a common standard for exit level exams, can considerably strengthen consistency in outcomes, implementation and assessment (Mummenthey et al., 2012). This will positively affect transferability of skills between workplaces, and thus the overall employability of learners.

In a context of few post-school opportunities, learnerships and apprenticeships are thus potentially significant routes to such critical vocational and occupational qualifications in South Africa, and the promise of future employment (Wildschut et al., 2012). They represent important alternative routes to enhance young peoples' transition to the labour market, and to meet the demand for scarce and critical skills. A 2008 review of SETAs showed that the skills development system suffers from weak reporting requirements, underdeveloped capacity, lack of effective management, and inadequate monitoring and evaluation that limit the ability of these institutions to serve as primary vehicles for skills development

(Marock et al., 2008). The foregoing shortcomings are indicative of management and evaluation weaknesses impacting the South African skills development system and they raise serious concerns about the quality of occupational learning, hence the present research which seeks to contribute to an effective solution.

3 Problem investigated

Prior to this research, no evidence was found which shows the existence of a valid and reliable measure for the effective management and evaluation of occupational learning programmes in the South African skills development context. Nevertheless, the following key problems which necessitate the development of a valid and reliable measure seem to exist in the South African occupational learning system:

(1) Challenges with regard to incoherent and inconsistent implementation of occupational learning continue to persist and this is evident in the literature (Grawitzky, 2007; Kraak, 2005; Mummenthey et al., 2012).

(2) Both the learnership and apprenticeship pathways are not operating optimally in South Africa (Kruss et al., 2012).

(3) The concept of 'Occupational Learning Programme' is still new in the South African skills development landscape, and SETAs and other stakeholders (skills development providers, employers, learners) are not clear regarding the elements and dimensions that comprise effective management and evaluation of occupational learning programmes (Tshilongamulenzhe, 2012b).

(4) There is no existing holistic and integrated management and evaluation model found in South Africa to date for occupational learning programmes (Tshilongamulenzhe, 2012b).

(5) There is no existing measure found in South Africa to date which assesses the effectiveness of management and evaluation practices pertaining to occupational learning programmes (Tshilongamulenzhe, 2012b).

The foregoing challenges coupled with the persistent skills shortage problem in the South African labour market, despite unprecedented policy interventions by government, have prompted the current research. Considering the enormous expectations for occupational learning programmes to provide an effective alternative towards addressing the skills deficit in South Africa, this research seems very important and profound. A valid and reliable measure will enhance management and evaluation practices pertaining to occupational learning programmes in South African workplaces and may potentially be used by SETAs and the Quality Council for Trades and Occupations (QCTO) to monitor the effectiveness of occupational learning programmes. It is envisaged that the application of the new measure will help

stakeholders in the skills development context in South Africa to manage and evaluate occupational learning programmes effectively in order to achieve the goals of the NSDS III and to improve the level of skills in the country.

4 Research objective

The objective of this research is to assess the psychometric properties of the Learning Programme Management and Evaluation (LPME) scale developed by Tshilongamulenzhe (2012b) as guided by the scale development framework of DeVellis (2012). The psychometric properties of the LPME scale were assessed in accordance with recommended practices (Gerbing and Anderson, 1988) and included assessments of some measures of content and construct validity and reliability. An examination of the psychometric properties of the LPME scale is necessary to ensure that the scale complies with the existing scientific conventions. This is also an important test to determine the rigour of the research process followed in the development of the new scale. The LPME scale seeks to ensure that occupational learning programmes are managed and evaluated effectively in the South African skills development context in order to achieve the goals of the NSDS III (2011-2016) (Tshilongamulenzhe et al., 2013). The newly developed scale was necessitated by the need for an integrated and coherent approach towards occupational learning programme management and evaluation with a view to effectively promote the alignment of skills development goals with the needs of the workplace, and with the broader growth and skills needs of the country's economy (DHET, 2010).

5 Focus of the article

In order to develop and assess the psychometric attributes of the LPME scale, this research was conducted in three phases, that is: scale development, scale refinement and scale validation. This article focuses on the first two phases of the research (scale development and scale refinement).

5.1 Phase 1: Scale development

Developing a valid and reliable scale is a process parallel to that aimed at constructing and testing a theory. As a result, scales go through a process of developing and testing. The aim is not only to develop a scale to allow theory testing but also to have a scale that is valid, reliable and reusable for other theories as well as for application purposes. Since no evidence was found in the literature which showed the existence of a valid and reliable scale to measure the effectiveness of learning programme management and evaluation practices in the South African skills development context, Tshilongamulenzhe (2012b) developed the Learning Programme Management and

Evaluation (LPME) scale as reported in this research and also as reported in Tshilongamulenzhe et al. (2013). The newly developed scale has to be subjected to a process of psychometric evaluation in order to ascertain its compliance with the established scientific conventions, hence this research. It is critical to assess the LPME scale for its validity, reliability, fit and dimensionality, and to determine the rigour of the research process followed in the development of this new scale. The process followed in the development of the new LPME scale is hereby outlined:

5.1.1 Item generation

In item generation, the primary concern is content validity, which may be viewed as the minimum psychometric requirement for measurement adequacy and the first step in the construct validation of a new scale (Schriesheim et al., 1993; Tshilongamulenzhe et al., 2013). Content validity must be built into the scale through the development of items (Tshilongamulenzhe et al., 2013). As such, any scale must adequately capture the specific domain of interest yet contain no extraneous content (DeVellis, 2003, 2012; Slavec and Drnovsek, 2012; Tshilongamulenzhe et al., 2013). There seems to be no generally accepted quantitative index of content validity of psychological scales; therefore judgement must be exercised in validating a scale (Stone, 1978; Tshilongamulenzhe et al., 2013). There are two basic approaches to item development that can be used during item generation (Fornaciari et al., 2005; Hinkin, 2009; Hunt, 1991). The first is deductive, sometimes called 'logical partitioning', or 'classification from above'. The second method is inductive, known also as 'grouping', or 'classification from below'.

Deductive scale development utilises a classification schema or typology prior to data collection (Hinkin, 2009; Hunt, 1991). This approach requires an understanding of the phenomenon to be investigated and a thorough review of the literature to develop the theoretical definition of the construct under scrutiny. The definition is then used as a guide for the development of items (Hinkin, 2009; Schwab, 1980). This approach can be used in two primary ways (Hinkin, 2009). First, researchers can derive items designed to tap into a previously defined theoretical universe. Second, researchers can develop conceptual definitions grounded in theory, but then utilise a sample of participants who are subject-matter experts to provide critical incidents that are subsequently used to develop items.

Conversely, the inductive approach is so labelled because there is often little theory involved at the outset as researchers attempt to identify constructs and generate scale items from individual responses (Hinkin, 2009; Hunt, 1991). According to Hinkin (2009), researchers usually develop scales inductively by asking a sample of participants to provide descriptions of their feelings about their organisations

or to describe some aspect of behaviour. Both deductive and inductively generated items may then be subjected to a sorting process that serves as a pre-test, permitting the deletion of items that are deemed to be conceptually inconsistent. To summarise, the generation of items is the most important part of developing sound scales (Hinkin, 2009; Worthington and Whittaker, 2006).

It is important to ensure that a clear link was established between items and their theoretical domain (Tshilongamulenzhe et al., 2013). In the current research, this was accomplished deductively by beginning with strong theoretical frameworks on skills development, the occupational learning system, training management and evaluation. The literature review allowed the researcher to identify and define 15 key constructs that were deemed relevant to a draft LPME scale. These constructs were scrutinised and grouped together under four defined elements which were later assembled into a theoretical framework for the effective management and evaluation of occupational learning programmes as proposed by Tshilongamulenzhe (2012b). This framework and its elements and constructs were thereafter used as a basis to set parameters and guide the item-generation process. A total list of 182 items was generated for a draft LPME scale.

5.1.2 Item development

At this stage of the process the researcher identifies a potential set of items for the construct or constructs under consideration. The elements and constructs identified in the previous stage, which constitute a theoretical framework as proposed by Tshilongamulenzhe (2012b), were used as parameters to sort the items rigorously in order to establish if they matched each of the elements and constructs of the theoretical framework. All 182 items were each matched to a relevant construct in the draft LPME scale with the guidance of existing theory.

The next task was the administration of these items to examine how well they confirmed expectations about the structure of the measure (Hinkin, 2009). This process included an assessment of the psychometric attributes such as validity, reliability, fit and dimensionality as reported later in this article. There has been considerable discussion regarding several important issues in measurement that have an effect on scale development (Hinkin, 2009). The first deals with the sample chosen, which should be representative of the population that the researcher will be studying in the future and to which results will be generalised. The sample chosen for the administration of items in this research was considered representative of the population as it contained all the key stakeholders in the occupational learning system (skills development providers, employers and learners/apprentices).

The second issue of concern was the use of negatively worded (reverse-scored) items. Such items may be employed primarily to ease response pattern bias (Hazlett-Stevens et al., 2004; Idaszak and Drasgow, 1987; Van Sonderen et al., 2013). The use of reverse-scored items has come under close scrutiny by a number of researchers. It has been shown to reduce the validity of questionnaire responses (Schriesheim and Hill, 1981) and may introduce systematic error to a scale (Jackson et al., 1993). Researchers have shown that reverse scoring may result in an artifactual response factor consisting of all negatively-worded items (Carlson et al., 2011; Harvey et al., 1985; Schmitt and Stults, 1985). In this study, no reverse-scored items were used.

The third issue concerns the number of items in a measure. Both adequate domain sampling and parsimony are important in order to obtain content and construct validity (Cronbach and Meehl, 1955). Total scale information is a function of the number of items in a scale, and scale lengths could affect the responses (Hazlett-Stevens et al., 2004; Roznowski, 1989; Van Sonderen et al., 2013; Worthington and Whittaker, 2006). Keeping a measure short is an effective means of minimising response biases (Schmitt and Stults, 1985; Schriesheim and Eisenbach, 1990; Worthington and Whittaker, 2006) but scales with too few items may lack content and construct validity, internal consistency and test-retest reliability (Nunnally, 1976; Kenny, 1979; Hinkin, 2009). Scales with too many items, on the other hand, can create problems such as respondent fatigue or response biases (Anastasi, 1976; Panther and Uys, 2008). Additional items also demand more time in both the development and administration of a measure (Carmines and Zeller, 1979; Hinkin, 2009). Adequate internal consistency reliabilities can be obtained with as few as three items (Cook et al., 1981; Hinkin, 2009) and adding items indefinitely progressively reduce and impact on scale reliability (Carmines and Zeller, 1979; Hinkin, 2009). In this research, a draft LPME scale consisted of an average of 12 items per construct, with a minimum range of three items to a maximum of 48 items for the 15 constructs identified by Tshilongamulenzhe (2012b).

With respect to the fourth issue, the scaling of items, it is important that the scale used should generate sufficient variance among participants for subsequent statistical analyses. In this research, the items were carefully worded, and the elements and constructs were properly defined to ensure that sufficient variance could be established among the participants.

The fifth issue is that of the sample size needed to conduct tests of statistical significance appropriately. The results of many multivariate techniques can be sample-specific, and increases in sample size may ameliorate this problem (Hinkin, 2009; Schwab, 1980). In simple terms, this means that, if powerful statistical tests and confidence in results are desired, the larger the sample, the better.

However, obtaining large samples could be very costly (Linacre, 1994; Stone, 1978; Verma and Burnett, 1996). As sample size increases, the likelihood of attaining statistical significance increases – it is important to note the difference between statistical and practical significance (Cohen, 1969). The current research had adequate sample ($n = 652$) to allow for the execution of a variety of statistical tests.

5.1.3 Item evaluation and refinement

At this stage, the review of the item pool begins. A team of about 20 expert reviewers was purposefully sampled and invited to review and assess the draft LPME measure. These experts were chosen in accordance with the three criteria prescribed by the Standard for Educational and Psychological Testing (American Psychological Association, 1985), that is, relevant training, experience and qualifications. The review team included expert academics and managers in the field of human resource development/training management. A group of about seven postgraduate students in the field of human resource development was also included. The total team was comprised of 27 participants. The review process focused on the face and content validation of the draft scale wherein the quality of the items was assessed in relation to the target population. Experts reviewed a pool of 182 items with instructions to assess the face, construct and content validity, to evaluate the relevance of the items to the constructs they proposed to measure, to assess the importance of the items, to assess the item difficulty level (easy, medium, difficult), and to judge items for clarity. Content Validity Ratio (CVR) proposed by Lawshe (1975) was used to estimate experts' perception of item relevance, importance and clarity. The CVR formula was applied after expert participants provided answers to three spectrums – 'item is relevant', 'item is important', and 'item is clear' – for each of the scale items (Tshilongamulenzhe et al., 2013). The goal was to obtain a reasonable number of items that would constitute the final draft LPME scale.

Item quality and content relevance for the final draft of the LPME scale were determined based on the strength of the literature and content experts' results and qualitative feedback. A decision to retain items for the final draft of the LPME scale was made based on the results of expert analysis, and on the acceptable qualitative feedback received regarding item clarity, difficulty, relevance and importance. The expert review results showed a clean ranking of each item in terms of clarity, difficulty, relevance, and importance (Tshilongamulenzhe et al., 2013). All the items were consistently ranked using CVR and the results ranged from an average CVR of .84 to 1 overall (Tshilongamulenzhe et al., 2013). However, as an average of less than 1 demonstrates that not all expert reviewers agree on the clarity, difficulty, relevance, and importance of some items, the researcher decided

to use a CVR cut-off point of .96 which is above the minimum of .90 used by Davis (1992). This was aimed to eliminate those items that were not clear, relevant and important to experts in the draft LPME scale in order to ensure that the instrument is valid as per the specified content domains as far as possible and to limit the variance error to less than 5% ($p < .05$).

According to Tshilongamulenzhe et al. (2013), the content expert results showed that 33 items had a CVR of 1, showing agreement across the board among experts; 76 items had a CVR ranging between .98 and .96. Four best-averaged items below a .96 cut-off point in two constructs were specially included in the final item pool to ensure that each construct had at least five items prior to exploratory factor analysis. Each pair of these four retained items had the highest CVR below the cut-off point (.93 and .94 respectively) in their theoretical constructs ('observation' and 'self-evaluation'). The revised draft LPME scale consisted of 113 items after the remaining 69 items below a CVR cut-off point of .96 had been eliminated (Tshilongamulenzhe et al., 2013). The revised draft LPME scale was then administered to the development sample for this research on a 6-point Likert-response format, ranging from (1) strongly agree to (6) strongly disagree. All items were classified into the appropriate construct and each construct had at least five items (Tshilongamulenzhe et al., 2013). As Benson and Clark (1982) state, a scale is considered to be content valid when the items adequately reflect the process and content dimensions of the specified aims of the scale as determined by expert opinion.

5.2 Phase 2: Scale refinement

This phase focused on the administration of the draft LPME scale on the development sample and the following are the materials and methods applied to achieve the objective of the research:

6 Methodology

6.1 Research approach

This phase followed a quantitative, non-experimental, cross-sectional survey design. Primary data collected from five Sector Education and Training Authorities (SETAs) and a human resource professional body in South Africa were used to achieve the objectives of this phase of the research.

6.2 Research participants

Participants in this study were 652 individuals drawn from six organisations: five SETAs and the South African Board for People Practices (SABPP), using a probabilistic simple random sampling technique. After permission had been obtained from the SETAs and the SABPP, a sample was extracted from the databases of

these organisations. These participants were diverse in their occupational status and included learning or training managers/employers, mentors/supervisors of learners/apprentices, skills development officers/providers, learning assessors/moderators as well as learners/apprentices. All sampled participants had to have some knowledge and understanding of the South African skills development context, including the new occupational learning system. The majority of the participants were young people trying to establish themselves in their careers. About 78.8% of the participants were aged below 35 years. Females constituted about 52.8% of the participants. In terms of educational achievement, 58.8% of the participants had acquired a senior certificate (matriculation/N3) as their highest qualification, with only 13.9% who had achieved a professional (4 years)/honours degree and higher. Regarding exposure to learning programmes, 86.6% of the participants were involved in learnerships, compared with just 13.4% who were involved in apprenticeships. In terms of current occupational commitments, over 65% of the participants constituted learners/apprentices, with 9% comprising employers/managers.

6.3 Measuring instrument

The revised draft LPME scale consisting of 113 items was used for data collection. This scale measured the elemental aspects outlined in the theoretical framework proposed by Tshilongamulenzhe (2012b).

6.4 Research procedure

The researcher wrote letters seeking permission to undertake this study to all 21 SETAs and the SABPP. Only the SABPP and five of the 21 SETAs gave permission for the research to be undertaken within

their jurisdictions. Once permission to undertake the research had been granted, the researcher started the process of planning for sampling and data collection with the respective organisations. The data collection process was carried out in the provinces of Gauteng, North West and Mpumalanga in South Africa.

6.5 Statistical analyses

Data for this study were analysed using the Statistical Package for Social Sciences (SPSS, Version 20) (IBM, 2011), Winsteps (Version 3.70.0) (Linacre, 2010) and Analysis of Moment Structures (AMOS) (Arbuckle, 2011). Exploratory factor analysis was executed using SPSS; Rasch analysis was executed using Winsteps; and structural equation modeling was executed using AMOS.

7 Results

7.1 Exploratory factor analysis

Exploratory factor analysis (EFA) was carried out in this research in order to reduce the number of items on the draft LPME scale into theoretically meaningful factors, to establish the underlying dimensions between these items and their constructs, and to provide evidence of construct validity. The EFA process began with the Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy and the Bartlett's Test of Sphericity in order to test the appropriateness of the data for factor analysis. It is evident in Table 1 that the KMO value of .960 was obtained, thus confirming the adequacy of sample for further statistical analysis. Specifically, the KMO value varies between 0 and 1, and values closer to 1 are better. The suggested minimum value that is acceptable for further analysis is .60 (Tabachnik and Fidell, 2001).

Table 1. KMO and Bartlett's test

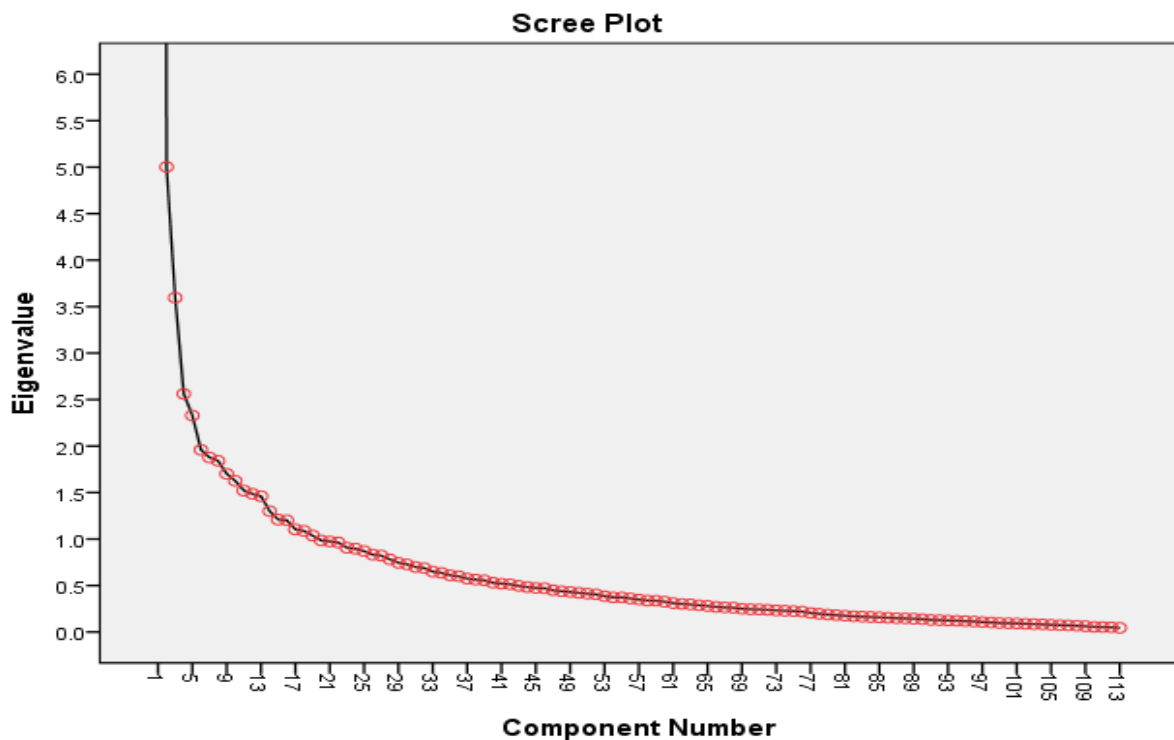
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.960
Bartlett's Test of Sphericity	Approx. Chi-Square	49316.106
	Df	6328
	Sig.	.000

Further to this, the Bartlett's Test of Sphericity was conducted, as depicted in Table 1, to test the null hypothesis that the correlation matrix is an identity matrix. An identity matrix is a matrix in which all the diagonal elements are 1 and off diagonal elements are 0 (Tshilongamulenzhe et al., 2013). The Bartlett's Test of Sphericity value was found to be statistically significant (df. 6328; $p < .05$) and rejects the null hypothesis that the correlation matrix is an identity matrix. The determinant of the correlation matrix between the factors was set to zero in this research due to orthogonal rotation restriction which suggests that

the factors cannot be correlated. Consequently, the results of these two tests show that data for this research can be subjected to Principal Components Analysis (PCA) and further statistical tests with confidence.

The next important task was to execute a PCA of all the 113 items of the draft LPME scale using varimax rotation in order to scientifically ascertain the constructs in which these items belong. Figure 1 presents the Scree Plot representation of all items of the draft LPME scale and the eigenvalue units of the matrix which are important in factor extraction.

Figure 1. Scree plot for factor retention



The factor extraction procedure followed an exploratory, iterative process using Kaiser's criterion and Catell's scree test in combination with the theory to specify the number of factors to be retained (Scheepers et al., 2008). The exploratory results are shown on the scree plot in Figure 1. Guided by existing theory and the quest to extract factors that could yield the most interpretable results with the least chance of random error, the researcher used the following criteria to determine the number of factors to be retained in the current research: a cut-off point of 1.45 eigenvalue units; items loading at .4 and higher; and a minimum of 4 items per factor. These criteria were also applied with success in other studies as reported by Tshilongamulenzhe et al. (2013). Based on the scree plot's representation in Figure 1, the first eleven factors were retained in this study as they met the stated criteria.

Subsequently, a total of 32 items which did not meet the criteria were eliminated. The remaining 81 items which constitute the final LPME scale as embedded in each of the eleven factors were retained, and each of these factors was considered a sub-scale of the LPME scale. A smallest sub-scale had 3 items while the largest had 16 items.

7.2 Separation indices, item fitness and reliability analysis

The results of this research include a summary of person/item separation indices, measure order, principal components analysis and reliability coefficients as shown in Table 2. Person and item separation and reliability of separation assess

instrument spread across the trait continuum (Green and Frantom, 2002). Separation measures the spread of both items and persons in standard error units. It can be thought of as the number of levels into which the sample of items and persons can be separated. For a measure to be useful, separation should exceed 1.0, with higher values of separation representing greater spread of items and persons along a continuum. Larger person/item separation indicates higher precision, meaning more distinct levels of function can be distinguished (Mallinson et al., 2004). For example, a person separation index ($G = 1.28$) for the Administrative Processes sub-scale as shown in Table 2 could reliably separate participants into at least two statistically distinct strata of persons (high ability and low ability persons). Similarly, the item separation index ($G = 3.59$) shows about five levels of item difficulty; very easy, easy, moderate, difficult and very difficult.

Lower values of separation indicate redundancy in the items and less variability of persons on the trait. If separation is 1.0 or below, then this may indicate that the items do not have sufficient breadth in position (Green and Frantom, 2002). In that case, it might be wise to reconsider what having less and more of the trait means in terms of items agreed or disagreed with, and on revision, add items that cover a broader range. An exception to this occurs if a measure is used to make dichotomous decisions.

Reliability of person separation was used in this research to demonstrate whether participants were being adequately separated by items along the continuum representing the construct, as well as provide an indication of replicability for person

placement across other items measuring the same construct. Equally important, the reliability of item separation was also examined to ensure that the measure adequately separates the people in terms of their ability.

The person separation indices for the LPME sub-scales as depicted in Table 2 ranged from .99 to 2.17, whereas the item separation indices range from .90 to 3.59. The expected mean value of both infit and outfit is 1 (Linacre and Wright, 1994; Planinic et al., 2010). Values < 1 suggest a lack of stochasticity in the data, potentially due to a violation of local independence – local independence means that, after controlling for

the latent trait, responses to items should be independent of each other (Fendrich et al., 2008). Values > 1 are indicative of excessive variability, which may signify a departure from unidimensionality. In this research, the average scale MNSQ Infit value was .99, meaning that there was a 1% deficiency in Rasch model predicted randomness in the data. Similarly, the average scale MNSQ Outfit value was .97, showing a 3% deficiency. However, the MNSQ Infit values for the LPME sub-scales ranged from .99 to 1.01; whereas the MNSQ Outfit values ranged from .97 to 1.04.

Table 2. Indices and fit statistics for the dimensions draft LPME scale

Sub-scale	No. of items	Separation Indices		Item Fit Statistics (Average Mean Scores)		Principal Components Analysis		Cronbach Alpha (α)
		Person Separation	Item Separation	MNSQ Infit	MNSQ Outfit	Variance Explained (%)	Unexplained Variance 1 st Contrast (Eigenvalues)	
Administrative Processes	4	1.28	3.59	1.00	1.00	45.2	1.9	.83
Environmental Scanning	6	1.06	2.02	.99	.99	45.8	1.4	.83
Monitoring and Evaluation	5	.99	3.01	.99	.99	46.4	1.7	.78
Observation and Problem Solving	6	1.50	2.59	.99	1.01	45.5	1.7	.88
Policy Awareness	8	1.74	1.58	1.00	.99	47.8	1.9	.89
Quality Assurance	4	1.08	2.69	.99	.97	44.7	1.6	.83
Stakeholder Inputs	17	2.17	2.98	1.00	1.01	51.8	1.7	.93
Strategic Leadership	4	1.17	2.27	.99	.96	56.9	1.6	.79
Learning Programme Design and Development	13	2.05	2.42	1.01	1.04	47.9	1.8	.92
Learning Programme Specifications	3	1.20	.90	.99	.78	61.1	1.6	.90
Occupational Competence	11	1.91	1.08	1.00	1.01	45.1	1.6	.92
Total Scale	81	1.46	2.28	.99	.97	48.9	1.6	.86

Outfit is based on the conventional averaged sum of squared standardised residuals, whereas infit is an information-weighted sum which gives more value to on-target observation (Planinic et al., 2010). A large infit value on a particular item indicates that some participants who had the ability to respond to difficult items did not respond in a way consistent with the model.

A large outfit value of an item indicates that persons who did not have the ability to respond to difficult items, responded in an unexpected way. For example, large outfit of an easy item means that some able persons have unexpectedly failed on that item. Larger outfit of a difficult item means that some persons of low ability have unexpectedly succeeded on this item. Large infit values are generally considered more problematic than larger outfit values. The results depicted in Table 2 do not show any evidence of excessive variability or deficiency regarding MNSQ Infit for all the LPME sub-scales.

The variance explained by the Rasch model as depicted in Table 2 is adequate with reasonable eigenvalues in the first contrast. To judge whether a

residual component adequately constitutes a separate dimension, the size of the first contrast eigenvalue (≥ 2) of unexplained variance must be attributable to this residual contrast. This suggests that all the LPME sub-scales are unidimensional as they have acceptable eigenvalue units (< 2) in the first contrast. The PCA of standardised residuals has an advantage over fit statistics in detecting departures from unidimensionality when (1) the level of common variance between components in multidimensional data increases and (2) there are approximately an equal number of items contributing to each component (Smith, 2004).

7.3 Structural Equation Modeling

The eleven-factor model of the LPME scale was tested by examining both overall model fit and the contribution of each indicator to the latent construct. The results are depicted in Tables 3 and 4. The factorial structure was tested using structural equation modeling to determine if the expected linear relationships existed between the latent construct and

its indicators of interest. According to Kelloway (1998), chi-square ratios of between 2 and 5 are regarded as indicative of good fit. Ratios less than 2 have been interpreted as indicating over-fitting (De Goede and Theron, 2010).

A Comparative Fit Index (CFI), Normed Fit Index (NFI) and Tucker Lewis Index (TLI) $\geq .95$ (Hu

and Bentler, 1999), and Root Mean Square Error of Approximation (RMSEA) of $\leq .06$ indicate good fit in the measurement model (Hu and Bentler, 1999). A Standardised Root Mean Square Residual (SRMR) of $\leq .80$ indicates a good model fit (Hu and Bentler, 1999).

Table 3. Summary of models

<i>Model</i>	<i>CMIN/DF</i>	<i>NFI</i>	<i>TLI</i>	<i>CFI</i>	<i>RMSEA</i>	<i>PCLOSE</i>	Δ <i>CMIN/DF</i>	<i>SRMR</i>
Criteria for a good fit	≤ 2	$\geq .95$	$\geq .95$	$\geq .95$	$\leq .06$	$\leq .05$	$\geq .01$	$\leq .08$
1. Initial model	5.659	.943	.929	.953	.090	.000	-	-
2. Revised model	1.847	.990	.987	.995	.038	.854	-3.812	-
3. Final model	3.363	.971	.964	.979	.064	.030	1.516	.0254

The results shown in Table 3 indicate that the initial hypothesized eleven-factor model did not fit the data well as its fit indices fell below the thresholds except for CFI (.95) and this model was revised. The revised model showed a good fit with the data ($df = 23$; $\chi^2/df = 1.84$; NFI = .99; TLI = .98; CFI = .99 and RMSEA = .03; PCLOSE $\leq .85$). However, evidence of significant variance of measurement error found during the analysis of the revised model data further directed the researcher to refine the revised model. The results of the final model are also presented in Table 3. It is clear that the χ^2 is significant at 127.81 ($df = 38$; $\chi^2/df = 3.36$). However, all other fit indices show that the final model fits the data perfectly (NFI = .97; TLI = .96; CFI = .97; RMSEA = .06; PCLOSE $\leq .03$; and SRMR = .02).

Table 4. Standardized regression weights for the initial and final models

<i>Sub-scale</i>		<i>Initial model</i>	<i>Final model</i>
Programme_DD	<--- OLP.	.865	.845
Policy_Awareness	<--- OLP.	.817	.810
Observation_PS	<--- OLP.	.815	.825
Quality_Assurance	<--- OLP.	.743	.720
Administrative_Processes	<--- OLP.	.764	.776
StakeholdersInputs	<--- OLP.	.859	.861
Strategic_Leadership	<--- OLP.	.533	.541
LearningPS	<--- OLP.	.740	.717
Monitoring_Evaluation	<--- OLP.	.675	.681
Occupational_Com	<--- OLP.	.800	.793
Enviromental_Scanning	<--- OLP.	.713	.692

Examining the data further, the researcher also analysed the standardized regression estimates in order to examine the extent of variance between the sub-scales of the LPME scale and the results are depicted in Table 4.

A standardized regression estimate (coefficient from an indicator variable to its construct) of $\geq .30$ indicates that a variable adequately contributes to the construct it was intended to measure (Kline, 2005). It is evident in Table 4 that the standardized regression

weights for the LPME sub-scales in both the initial and final model are adequate to support model fit as they range between .53 to .86, and .54 to .86 respectively. All the estimates are positive and statistically significant, and they surpass the $\geq .30$ value suggested by Kline (2005).

An analysis of all individual parameters in the final model shows that all eleven sub-scales of the LPME scale were significant predictors of an effective occupational learning programme (critical ratios (CR) were statistically significant and ranged from 13.790 to 25.528) and supported the validity, reliability and dimensionality of the LPME scale and its sub-scales.

8 Discussion

The focus of this research was to assess and report the psychometric attributes of the LPME scale developed by Tshilongamulenzhe (2012b) who was guided by Scheepers et al. (2008) proposition that a valid measurement scale must be viewed as a pre-condition for the successful study of phenomena in business and science. According to Boshoff (2009) and Terblanche and Boshoff (2006), reliability assessment and validity checks have improved in recent years due to the availability of statistical procedures such as exploratory and confirmatory factor analysis which provide additional evidence of construct validity. The psychometric attributes of the LPME scale were examined in accordance with the recommended practices as suggested by Gerbing and Anderson (1988), and these included assessment of measures of scale validity, reliability, fit and dimensionality.

Construct validity was measured when the content of each item of the LPME scale was rigorously matched with the constructs of this research during the item development phase. Content validity was measured when the items of the LPME scale were checked for consistency with the definition of the elements and constructs of the theoretical framework proposed by Tshilongamulenzhe (2012b). Discriminant validity was measured during exploratory factor analysis. As suggested by Tabachnick and Fidel (2001), the results of the KMO

Measure of Sampling Adequacy and Bartlett's Test of Sphericity provided indications that the data of this research were suitable for factor analysis. Catell's scree test, as depicted in Figure 1, identified the factors that have a substantial amount of common variance before the inflection point and these were retained as they contribute the most to the explanation of the variance in the data set. These rigorous statistical procedures executed in this research and the findings show that the LPME scale is valid and reliable, and complies with the psychometric expectations.

The validity and reliability claim on a new scale depends on the methodology used to arrive at the verdict. According to De Goede and Theron (2010), methodology is meant to serve the epistemic ideal of science and if the methodology used is not made explicit, evaluation of the researcher's conclusions become difficult. Under such circumstances, the rationality of science suffers, as does ultimately the epistemic ideal of science (Babbie and Mouton, 2004). A comprehensive account of the methodology applied in this research was succinctly described. The development of the LPME scale was guided by the framework of DeVellis (2012) and all stages of this framework were successfully applied, which further enhances the validity and reliability of the scale. The reliability of the LPME scale and its sub-scales was measured using Cronbach's alpha coefficient and the findings show an overall coefficient of .86 for the scale, while those of the sub-scales ranged from .78 to .93. These coefficients are above the cut-off point of $\geq .70$ which is considered acceptable (Kline, 2005). The reliability coefficient results are depicted in Table 2. A further measurement of reliability was conducted using the Rasch model through an assessment of separation indices, item fit statistics and principal component analysis for unidimensionality. The results show an average person separation index of 1.46 and an item separation index of 2.28. The average item MNSQ infit was .99 while the MNSQ outfit was .97. The average variance explained by the LPME scale relative to the Rasch model was 48.9% with 1.6 eigenvalue units. These results are depicted in Table 2 and they support the validity and reliability of the LPME scale.

De Goede and Theron (2010) suggest that, in order to come to valid and credible conclusions about the ability of the structural model to explain the pattern of covariance among the indicator variables, evidence is required that the manifest indicators are indeed valid and reliable measures of the latent variables they are linked to. The results of structural equation modelling computed in this research as depicted in Tables 3 and 4 indicate a good model fit for the LPME scale (NFI = .97; CFI = .97; RMSEA = .06 and SRMR = .02). The results show that all 11 dimensions were significant predictors of occupational learning programmes, and this further supports the validity and reliability of the LPME scale. Overall, the

findings of this research led to a conclusion that the LPME scale is valid, reliable and unidimensional, and can be applied with confidence in the South African skills development context.

9 Limitation, implications and recommendations

However, irrespective of the contributions made by this research, a limitation is that cross-validation of the LPME scale on a different sample has not been done yet. Therefore, the findings are based on data obtained from the original development sample of this research.

The findings of this research have the following implications for practice within the South African skills development context:

- The LPME scale and its sub-scales should be seen as a window of opportunity for future research initiatives focusing on the management and evaluation of occupational learning programmes. The sub-scales of the LPME scale can be applied autonomously.

- Scholars in the sub-field of training management/human resource development should find it possible to use the findings of this research as a baseline input to further critique and refine the LPME scale and its sub-scales.

- The LPME scale and its sub-scales should enable relevant occupational learning stakeholders to diagnose weaknesses in the system so that appropriate remedial action can be taken using a scientific tool.

- SETAs, skills development providers, and employers should use the LPME scale and its sub-scales in their task of managing, monitoring and evaluating the feasibility and success of learning programme implementation in their respective contexts.

The findings of this study provide direction for future enquiry by suggesting a cross-validation study and an action research whereby the newly developed LPME scale and its sub-scales could be applied and evaluated in practice. To this end, reliability is regarded as a necessary condition for validity.

References

1. American Psychological Association, (1985), *Standards for Educational and Psychological Testing*, American Psychological Association, Washington, DC.
2. Anastasi, A, (1976), *Psychological testing*, 4th edn, Macmillan, New York.
3. Arbuckle, J. L., (2011), *IBM SPSS Amos 20 User Guide*, Amos Development Corporation, New York.
4. Babbie, E. and Mouton, J, (2004), *The Practice of Social Research*, 4th edn, Oxford University Press, Cape Town.
5. Benson, J. and Clark, F, (1982), "A Guide for Instrument Development and Validation", *The American Journal of Occupational Therapy*, Vol. 36, pp. 789-800.
6. Boshoff C, (2009), "A Psychometric Assessment of an Instrument to Measure a Service Firm's Customer-

- Based Corporate Reputation". *South African Journal of Business Management*, Vol. 40, 35-44.
7. Carlson, M. Wilcox, R. Chou, C. Chang, M. Yang, F. Blanchard, J. Marterella, A. Kuo, A. and Clark, F. (2011), "Psychometric Properties of Reverse-Scored Items on the CES-D in a Sample of Ethnically Diverse Older Adults". *Psychological Assessment*, Vol. 23 No. 2, pp. 558-562.
 8. Carmines, E. G. and Zeller, R. A. (1979), *Reliability and validity assessment*. Sage, Beverly Hills.
 9. Coetzee, M. Botha, J. Kiley, J. Truman, K. and Tshilongamulenzhe, M.C. (2012), *Practising Training and Development in South Africa*, 2nd edn, Juta & Company Ltd, Cape Town.
 10. Cohen, J. (1969), *Statistical power analysis for the behavioural sciences*, Academic Press, New York.
 11. Cook, J.D. Hepworth, S.J. Wall, T.D. and Warr, P.B. (1981), *The experience of work*, Academic Press, San Diego.
 12. Cronbach, L.J. and Meehl, P.C. (1955), "Construct validity in psychological tests", *Psychological Bulletin*, Vol. 52, pp. 281-302.
 13. Davis, L. L. (1992), "Instrument review: Getting the most from a panel of experts", *Applied Nursing Research*, Vol. 5, pp. 194-197.
 14. De Goede, J. and Theron, C. (2010), "An investigation into the internal structure of the learning potential construct as measured by the APIL-B Test Battery", *Management Dynamics*, Vol. 19, pp. 30-53.
 15. DeVellis, R.F. (2012), *Scale Development: Theory and Application*, 3rd edn, Sage Publications, Thousand Oaks, CA.
 16. DeVellis, R.F. (2003), *Scale development*, 2nd edn, Sage Publications, Thousand Oaks, CA.
 17. DHET, (2014), *National List of Occupations in High Demand*, Department of Higher Education and Training, Pretoria.
 18. DHET, (2010), *Framework for the National Skills Development Strategy 2011/12 – 2015/16: First Draft for Consultation, 29th April*, Department of Higher Education and Training, Pretoria.
 19. Du Toit, R. (2012), *The NSF as a mechanism to address skills development of the unemployed in South Africa*, HSRC, Pretoria.
 20. Erasmus, J. and Breier, M. (2009), *Skills shortages in South Africa: Case studies of key professions*, HSRC Press, Pretoria.
 21. Fendrich, M. Smith, E.V Jr. Pollack, L.M. and Mackesy-Amiti, M.E. (2008), "Measuring Sexual Risk for HIV: A Rasch Scaling Approach", *Archives of Sexual Behaviour*, Vol. 38, pp. 922-935.
 22. Fornaciari, C.J. Sherlock, J.J. Ritchie, W.J. and Dean, K.L. (2005), "Scale Development Practices in the Measurement of Spirituality", *International Journal of Organizational Analysis*, Vol. 13 No. 1, pp. 28-49.
 23. Gerbing, D.W. and Anderson, J.C. (1988), "An Updated Paradigm for Scale Development Incorporating Unidimensionality and its Assessment", *Journal of Marketing Research*, Vol. 25, pp. 186-192.
 24. Grawitzky, R. (2007), *SETAs - A vehicle for the skills evolution?*, Development Policy Research Unit (DPRU), University of Cape Town.
 25. Green, K.E. and Frantom, C.G. (2002), *Survey Development and Validation with the Rasch model*. A paper presented at the International Conference on Questionnaire Development, Evaluation, and Testing, Charleston, SC, November 14-17.
 26. Harvey, R. Billings, R. and Nilan, K. (1985), "Confirmatory factor analysis of the Job Diagnostic Survey: good news and bad news", *Journal of Applied Psychology*, Vol. 70, pp. 461-468.
 27. Hinkin, T.R. (2009), *Scale Development Principles and Practices*, in Swanson, R.A and Holton III, E.F (Eds.), *Research in Organisations: Foundations and Methods of Inquiry*, Berret-Koehler Publisher, San Francisco, CA, pp. 161-180.
 28. Hu, L.T. and Bentler, P.M. (1999), "Cut-Off Criteria for Fit Indices in Covariance Structure Analysis: Conventional Criteria versus New Alternatives", *Structural Equation Modeling*, Vol. 6, pp. 1-55.
 29. Hunt, S.D. (1991), *Modern Marketing Theory*, South-Western Publishing, Cincinnati.
 30. IBM, (2011), *IBM SPSS Statistics 20 Brief Guide*, IBM Corporation, New York.
 31. Idaszak, J.R. and Drasgow, F. (1987), "A revision of the Job Diagnostic Survey: Elimination of a measurement artefact". *Journal of Applied Psychology*, Vol. 72, pp. 69-74.
 32. Jackson, P.R. Wall, T.D. Martin, R. and Davids, K. (1993), "New measures of job control, cognitive demand, and production responsibility", *Journal of Applied Psychology*, Vol. 78, pp. 753-762.
 33. Kelloway, E.K. (1998), *Using LISREL for Structural Equation Modelling: A Researcher's Guide*, Sage Publications, New York.
 34. Kenny, D.A. (1979), *Correlation and causality*, Wiley, New York.
 35. Kline, R.B. (2005), *Principles and Practice of Structural Equation Modelling*, 2nd edn, Guilford Press, New York.
 36. Kraak, A. (2005), "Human resources development and the skills crisis in South Africa: the need for a multi-prolonged strategy", *Journal of Education and Work*, Vol. 18 No. 1, pp. 57-83.
 37. Kruss, G. Wildschut, A. Janse Van Rensburg, D. Visser, M. Haupt, G. and Roodt, J. (2012), *Developing skills and capabilities through the learnership and apprenticeship pathway systems. Project synthesis report. Assessing the impact of learnerships and apprenticeships under NSDSII*, HSRC, Pretoria.
 38. Lawshe, C.H. (1975), "A quantitative approach to content validity", *Personnel Psychology*, Vol. 28, pp. 563-575.
 39. Linacre, J.M. (2010), *Winsteps® (Version 3.70.0)* [Computer Software]. Winsteps.com, Beaverton, Oregon.
 40. Linacre, J.M. (1994), "Sample Size and Item Calibrations Stability", *Rasch Measurement Transaction*, Vol. 7 No. 4, pp. 328.
 41. Linacre, J.M. and Wright, B.D. (1994), "Reasonable mean square fit values", *Rasch Measurement Transactions*, Vol. 8, p. 370.
 42. Lundall, P. (2003), *Sector Training and Development Authorities and the Delivery of Training: Preliminary Remarks on the New Skills Dispensation in South Africa*, Development Policy Research Unit, University of Cape Town.
 43. Mallinson, T. Stelmack, J. and Velozo, C. (2004), "A Comparison of the Separation Ratio and Coefficient a in the Creation of Minimum Item Sets", *Medical Care*, Vol. 42, pp. 117-124.
 44. Marock, C. Harrison-Train, C. Soobrayan, B. and Gunthorpe, J. (2008), *SETA review*, A paper presented during the Conference on the Regulatory Environment

- and its Impact on the Nature and Level of Economic Growth and Development in South Africa, Development Policy Research Unit, University of Cape Town.
45. Mummenthey, C. Wildschut, A. and Kruss, G. (2012), *Assessing the impact of learnerships and apprenticeships under NSDSII: Three case studies: MERSETA, FASSET & HWSETA*, HSRC, Pretoria.
 46. Nunnally, J.C. (1976), *Psychometric theory*, 2nd edn, McGraw-Hill, New York.
 47. Panther, S. and Uys, C.S. (2008), "Using scale reduction technique for improved quality of survey information", *South African Journal of Information Management*, Vol. 10 No. 3, pp. 1-14.
 48. Planinic, M. Ivanjek, L. and Susac, A. (2010), "Rasch Model Based Analysis of the Force Concept Inventory", *Physical Review Special Topics – Physical Education Research*, Vol. 6, pp. 1-11.
 49. Republic of South Africa, (2008), *Skills Development Amendment Act, Act No. 37 of 2008*, Government Printers, Pretoria.
 50. Roznowski, M. (1989), "Examination of the measurement properties of the Job Descriptive Index with experimental items", *Journal of Applied Psychology*, Vol. 74, pp. 805-814.
 51. Scheepers, M.J. Bloom, J.Z. and Hough, J. (2008), "The development of an instrument to assess the enacted environment for corporate entrepreneurship in South Africa", *Management Dynamics*, Vol. 17, pp. 1-17.
 52. Schmitt, N.W. and Stults, D.M. (1985), "Factors defined by negatively keyed items: The results of careless respondents?", *Applied Psychological Measurement*, Vol. 9, pp. 367-373.
 53. Schriesheim, C.A. and Eisenbach, R.J. (1990), *Item wording effects on exploratory factor-analytic results: An experimental investigation*, Proceedings of the 1990 Southern Management Association annual meetings, Washington D.C.
 54. Schriesheim, C.A. and Hill, K. (1981), "Controlling acquiescence response bias by item reversal: The effect on questionnaire validity", *Educational and psychological measurement*, Vol. 41, pp. 1101-1114.
 55. Schriesheim, C.A. Powers, K.J. Scandura, T.A. Gardiner, C.C. and Lankau, M.J. (1993), "Improving Construct Measurement in Management Research: Comments and a Quantitative Approach for Assessing the Theoretical Content Adequacy of Paper-and-Pencil Survey-Type Instruments", *Journal of Management*, Vol. 19, pp. 385-417.
 56. Schwab, D.P. (1980), Construct Validity in Organisation Behaviour, in Staw, B.M and Cummings, L.L (Eds.), *Research in Organisational Behaviour*, JAI Press, Greenwich, CT, pp. 3-43.
 57. Sebusi, I.E. (2007), *An economic analysis of the skills shortage problem in South Africa* (Unpublished Master's Dissertation), University of Johannesburg, Johannesburg.
 58. Skinner, D. Saunders, M.N.K. and Beresford, R. (2004), "Towards a shared understanding of skill shortages: differing perceptions of training and development needs" *Journal of Education and Training*, Vol. 46 No. 4, pp. 182-183.
 59. Slavec, A. and Drnovsek, M. (2012), "A Perspective on Scale Development in Entrepreneurship Research", *Economic and Business Review*, Vol. 14 No. 1, pp. 39-62.
 60. Smith, E.V Jr, (2004), Evidence for the Reliability of Measures and Validity of Measure Interpretation: A Rasch Measurement Perspective, in Smith, E.V Jr and Smith, R.M (Eds.), *Introduction to Rasch Measurement*, JAM Press, Maple Grove, Minnesota, pp.93-122.
 61. SRI, (2008), *Challenges facing SA economy: Skills biggest constraint to economic growth*, Solidarity Research Institute, Johannesburg.
 62. Stone, E. (1978), *Research Methods in Organisational Behaviour*, Scott, Foresman, Glenview, IL.
 63. Tabachnick, B.G. and Fidell, L.S. (2001), *Using Multivariate Statistics*, 4th edn, Allyn & Bacon, Needham Heights, MA.
 64. Terblanche, N.S. and Boshoff, C. (2006), "A Generic Instrument to Measure Customer Satisfaction with the Controllable Elements of the In-Store Shopping Experience", *South African Journal of Business Management*, Vol. 37, pp. 1-15.
 65. Townsend, V. (2006), *The skills shortage: Is history repeating itself?*, ICMEEESA, South Africa.
 66. Tshilongamulenzhe, M.C. (2012a), "Enunciating the skills development challenge facing South Africa", *GSTF Journal of Law and Social Sciences*, Vol. 1 No. 1, 30-35.
 67. Tshilongamulenzhe, M.C. (2012b), *An Integrated Learning Programme Management and Evaluation Model for the South African Skills Development Context* (Unpublished Doctorate Thesis), University of South Africa, Pretoria.
 68. Tshilongamulenzhe, M.C. Coetzee, M. and Masenge, A. (2013), "Development of the Learning Programme Management and Evaluation Scale for the South African Skills Development Context", *South African Journal of Industrial Psychology*, Vol. 39 No.1, 131-144.
 69. Van Rooyen, M. (2009), *New learning programmes leading to new qualifications*, Cutting Edge, Johannesburg.
 70. Van Sonderen, E. Sanderman, R. and Coyne, J.C. (2013), "Ineffectiveness of Reverse Wording of Questionnaire Items: Let's Learn from Cows in the Rain", *Public Library of Science*, Vol. 8 No. 7, pp. 1-22.
 71. Verma, S. and Burnett, M.F. (1996), "Cutting Evaluation Costs by Reducing Sample Size", *Journal of Extension*, Vol. 43 No. 1, retrieved on 08 December 2014 from <http://www.joe.org/joe/1996february/a2.php>
 72. Visser, M. and Kruss, G. (2009), "Learnerships and skills development in South Africa: A shift to prioritise the young unemployed", *Journal of Vocational Education and Training*, Vol. 61 No. 3, pp. 357-374.
 73. Wildschut, A. Kruss, G. Janse Van Rensburg, D. Haupt, G. and Visser, M. (2012), *Learnerships and Apprenticeships Survey 2010 Technical Report: Identifying Transitions and Trajectories through the Learnership and Apprenticeship Systems*, HSRC, Pretoria.
 74. Worthington, R.L. and Whittaker, T.A. (2006), "Scale Development Research: A Content Analysis and Recommendations for Best Practices", *The Counseling Psychologist*, Vol. 34 No. 6, pp. 806-838.