

# DEVELOPMENT OF A QUESTIONNAIRE FOR ASSESSING WORK UNIT PERFORMANCE

H H SPANGENBERG

*Centre for Leadership Studies*

*Graduate School of Business*

*University of Stellenbosch*

C THERON

*Department of Industrial Psychology*

*University of Stellenbosch*

## ABSTRACT

The aim of this study was to develop a generic work unit performance measure that can be utilized in private, public and non-profit work units. Development of such a questionnaire, called the Performance Index (PI) comprised three steps, namely deciding on a baseline structure for the model; verifying the model and dimensions through subject expert feedback; and consolidating the model and refining the questionnaire. The sample consisted of 60 units rated on a 360 basis by 257 respondents. Item and dimensionality analyses, followed by confirmatory factor analysis utilizing LISREL produced acceptable model fit. Overall, results provided reasonable psychometric support for the Performance Index.

## OPSOMMING

Die doel van die studie was om 'n generiese werkeenheid-prestasiemeting te ontwikkel wat in privaat, openbare en nie-winsgerigte organisasies gebruik kan word. Die ontwikkeling van die vraelys, die Prestasie-Indeks (PI) het drie stappe behels, naamlik om te besluit op 'n onderliggende struktuur vir die model; om die model en dimensies deur middel van ekspert-beoordeling te verifieer; en om die model te konsolideer en die vraelys af te rond. Die steekproef het bestaan uit 60 eenhede wat op 'n 360-basis deur 257 respondente beoordeel is. Item- en dimensionaliteitsontledings, gevolg deur Lisrel-gedrewe bevestigende faktorontleding het gedui op aanvaarbare modelpassing. In geheel beslou het resultate redelike psigometrieë ondersteuning aan die Prestasie-Indeks verleen.

The need for a measure of work unit performance that could serve as a criterion variable first arose when the authors wanted to validate a Performance Management Audit Questionnaire, and more recently when the Centre of Leadership Studies (Southern Africa) decided to develop a leadership questionnaire that accommodates the challenges facing South African leaders. No satisfactory criterion measure of unit performance seemed to exist that encompassed all the unit performance dimensions for which the unit leader could be held accountable. In addition to utilizing a measure of work unit performance for validation research purposes, such an assessment instrument could comprise a major part of overall organizational diagnosis as used in organization development and change programmes.

Only two psychometric measures of organizational performance were traced in the research literature that were applicable to this study, namely Nicholson and Brenner's (1994) dimensions of perceived organizational performance that are based on a systems model, and the Unit Performance Questionnaire (UPQ) (Cockerill, Schroder & Hunt, 1993; Spangenberg, Schroder, Duvenage & Theron, 1999). The Nicholson and Brenner (1994) model comprises four elements, namely wealth, e.g. resources, profits, assets, investments and skilled staff; markets e.g. sources of supply, customer base and market position; adaptability; and climate. Expected future growth served as an overall index of future expected performance. The rationale for including the four criteria was as follows: First, there was agreement in the literature on the multi-dimensionality of performance and effectiveness (Cameron, 1986; Cameron & Whetton, 1983; Lewin & Minton, 1986; Quinn & Rohrbaugh, 1983). Second, experimentation with the Miles and Snow (1978) model indicated the need to distinguish between internal strength (wealth), external capability (markets), and ability to manage change (adaptability). Third, their experimentation suggested that

climate as a global perception of the psychological environment was a prerequisite for integrated functioning among employees. The four dimensions of effectiveness, i. e. wealth, markets, adaptability, and climate are measured by seventeen items. Respondents are asked to rate their organization against its main external competitors, using the following item-response scales: "much weaker/weaker/no different/stronger/much stronger" and "don't know". Future growth is rated in terms of growth prospects for the following five years in the respondent's area of the organization. Item-response scales were: "major decline/some decline/no change/some growth/major growth".

The UPQ is a 360° instrument adapted from Nicholson and Brenner's (1994) performance measure. It measures four dimensions of work unit performance, namely output, climate, adaptability and resource input. Classical factor analysis of the UPQ indicated that the 18 items comprising the questionnaire loaded on the four factors in accordance with the Nicholson and Brenner model (Cockerill et al., 1993). The four factors explained 59.4% of the variance. Cronbach alpha reliability coefficients for the subscales were all above 0,80, with the exception of the resources subscale, which returned a value of 0,72 (n = 1649). In a recent study, Spangenberg et al. (1999) used the UPQ to validate the Schroder high-performance leadership competencies measured by means of a multi-rater competence utilization questionnaire (Schroder & Cockerill, 1992). The reliability of three subscale measures derived from the combined 360° item ratings was satisfactory (Cronbach alpha coefficients were 0,79 and higher, n = 186). The reliability of the resources-scale, however, again provided reason for concern ( $\alpha = 0,60$ ). The reliability of the unweighted linearly combined overall unit performance measures was calculated in accordance with the formula suggested by Nunnally (1978), and was high (0,95).

Neither the UPQ nor the Nicholson and Brenner (1994) performance measure covers the unit performance domain comprehensively enough to successfully serve the purpose of

a work unit criterion measure. These shortcomings indicate the need for a generic, standardized work unit performance measure that can readily be used in private, public and non-profit work units or departments. The reason for developing a generic rather than a unit-specific measure stems from the need to validate leadership and Performance Management models across a spectrum of units and organizations. The objective of this study is, therefore, the development of such an instrument. The concepts work unit and department are interpreted to have the same meaning, and in this text the term work unit will be used. A work unit is defined as a permanent or semi-permanent organizational entity, nested in a public, private or not-for-profit organization with specific, identifiable and measurable performance goals for which it is held accountable by higher management structures. The size of a work unit may vary from small, i.e. a leader and at least three followers to large, comprising a large staff complement.

There exists a possibility that the concepts of team and work unit might be confused. For the purpose of this study, the two concepts are not considered synonymous, although they are clearly related. Given the complex nature of the business environment, and consequently the need for adaptability, we argue that an appropriate degree of teamwork (synergistic interaction) is required for effective work unit performance. The relevance of team and teamwork to work unit functioning is acknowledged in the proposed model through the inclusion of synergistic interaction, reflected by core people processes and climate.

The literature review that follows focuses on issues that are central to the development of a work unit performance measurement questionnaire, namely organizational effectiveness that underlies the measurement of organizational performance, the time-dimension model of organizational effectiveness, and the validity of non-financial performance measures. Since a work unit forms an integral part of the organisation, organizational effectiveness parameters apply to it as well. An organization is, in effect, an aggregate of work units guided by a unifying mission and vision and, therefore, in itself also a work unit. The work unit is naturally exposed to the same contextual forces, internal and external as the rest of the organization.

#### **Organizational effectiveness**

Two prominent models of organizational effectiveness described in the literature, are the goal approach and the systems approach (Denison, 1990; Eccles, 1991; Miles, 1980). The usefulness of the goal approach lies in the specificity of the performance information it provides. On the other hand, the organizational and content specificity of these measures prohibits comparison of effectiveness between different functional units. This problem applies to both intra-organizational unit comparisons in the case of large organizations, as well as to comparisons across different organizations. Another problem is that goal achievement is not easily measurable in organizations that do not produce tangible results (Gibson, Ivancevich & Donnelly, 1991). Therefore, due to difficulties in circumscribing organizational goals, this model is considered unsatisfactory for measuring unit performance (Miles, 1980; Mohr, 1983).

Weaknesses of the goal model led to the development of systems models of organizational effectiveness, which focus on the means to achieve the objectives of organizations rather than on the ends themselves (Miles, 1980). The systems approach focuses on system boundaries, differentiation and integration of subsystems that are part of the total system, input-transformation-output processes, boundary transactions and system maintenance processes (Seashore, 1983). There is a contention that the construct of organizational effectiveness should be described in terms of

all the attributes of the system that significantly impact on its processes of adaptation, maintenance and transformation (Quinn & Rohrbaugh, 1983). The main outcomes of the systems model are survival, growth, and stability or decline (Denison, 1990).

The systemic, integrated and adaptive approach to organizational effectiveness reflected by the systems model calls for a diverse set of performance measures. In addition, since the beginning of the 1990s there has been dissatisfaction with the near exclusive use of financial performance measures (Eccles, 1991). The main causes of this dissatisfaction with financial performance measures are perceived limitations in traditional accounting-based measures, namely:

1. Diversity in products, markets and business units puts a big strain on accounting theories and rules developed for smaller, less complex organizations.
2. Traditional accounting measures often fail to support investments in new technologies and markets that are essential for successful performance in global markets. It seems that these measures do not capture key business changes until it is too late.
3. Such criticisms reinforce concern about harmful management practices such as short-term thinking on competitiveness of companies. There is also concern about rewarding short-term behavior (Brancato, 1995; Eccles, 1991; Fisher, 1995; Ittner & Larcker (1998a).

A logical consequence was, therefore, the use of both financial and non-financial measures. Most economic theories that analyze performance measures indicate that performance measurement and reward systems should include both financial and non-financial measures (Ittner & Larcker, 1998a). In addition to shortcomings of financial measures described above, the need for non-financial measures has been stimulated by major changes in the external business environment such as:

1. A shift from tangible assets to intangible assets. Towards the 1990s intangible assets became the major source for competitive advantage.
2. Competitive pressure. By incorporating forward-looking non-financial measures such as customer satisfaction, employee satisfaction and defect rates, a wider set of measures now captures factors that lead to the creation of value for the organization (Ittner & Larcker, 1998a).
3. The quality movement (TQM) which requires greater emphasis on customer requirements and customer satisfaction, leading to greater emphasis on non-financial measures such as complaints, satisfaction and retention. Some organizations utilize quality improvement as a major intervention to upgrade their measurement process (Brancato, 1995).

#### **A time-dimension model of organizational effectiveness**

A major factor that relates to the systems model of effectiveness is time. Considering that the organization is part of a larger system, namely the environment, through time the organization acquires processes and returns resources to the environment. The ultimate criterion of organizational effectiveness is sustainability in the environment. Survival of the organization is, therefore, the long-term criterion of effectiveness. The time-dimension model defines effectiveness criteria over the short term, medium term and long term (Gibson et al., 1991, p37). In view of the expressed need for non-financial measures to facilitate the creation of value for the organization, the time-dimension model of effectiveness (Gibson et al., 1991) was extended to include these measures as well. Furthermore, the adapted model distinguishes between financial and non-financial measures. For the sake of clarity the term production was replaced by outputs (see Figure 1).

Time					
Short-term		Medium-term		Long-term	
Financial	Non-financial	Financial	Non-financial	Financial	Non-financial
<b>Outputs (quantity &amp; quality)</b>	<b>Outputs (quantity &amp; quality)</b>	<b>Capacity/Development</b>	<b>Capacity/development</b>	<b>Market standing</b>	<b>Market standing</b>
1. Profit	1. Units delivered	1. Investment	1. Competent staff	1. Market share	1. Reputation
2. Economic value added	<b>Efficiency</b>	2. Financial resources	2. Physical assets	<b>Survival &amp; future growth</b>	2. Competitiveness
3. Revenue growth	1. Defect rates		3. Materials supply	1. Market share	<b>Survival &amp; future growth</b>
4. Sales	2. On-time delivery		<b>Adaptability</b>	2. Capital investment	1. Staff levels
5. Market share	3. Machine utilization		1. to external change	3. Profits	2. Acquisitions
<b>Efficiency</b>	4. Scrappages		2. to internal change		
1. Return on investment	<b>Employee satisfaction</b>		3. management plans		
2. Return on assets	1. Turnover		<b>Climate</b>		
2. Cash flow return on investment	2. Absenteeism		1. Teamwork		
	3. Tardiness		2. Cohesion		
	4. Grievances		3. Shared values		
	<b>Customer satisfaction</b>				

**Figure 1: An adapted time-dimension model of financial and non-financial performance measures**

*Short-term* measures of performance have been discussed comprehensively in the literature and do not need further deliberation here. However, some major studies are cited: 1. Outputs (quantity and quality). Profits and economic value measures (Balkcom, Ittner & Larcker, 1997), quality (Youndt, Snell, Dean & Lepak, 1996). 2. Efficiency. Financial efficiency (Rucci, Kirn & Quinn, 1998), production efficiency (Banker, Field, Schroeder & Sinha, 1996). 3. Satisfaction. Employee satisfaction (Huang, 1997), customer satisfaction (Delaney & Huselid, 1996). *Medium-term* non-financial performance measures are considered as important value creators in organizations. Three growth-orientated measures will be discussed, namely development, adaptability, and climate. In order to meet future environmental demands, organizations have to invest resources for development wisely. This includes continued investment in production capacity, technology, and in expanding the capabilities of managerial and non-managerial staff. Gibson et al. (1991) wisely argued that future oriented investment of resources may reduce production and efficiency in the short term, but, if properly managed, development efforts often provide the key to survival.

It is important for the organization to adapt to changes in its external and internal environments. Nicholson and Brenner (1994) tested perceptual measures of inter-relatedness amongst the elements of their systems model. The main finding of their study was that adaptability emerged as the core of organizational effectiveness. This applied either when adaptability was directly associated with other outcomes or when it mediated them. This finding is considered as consistent with the view that mastery of uncertainty (adaptability) is a survival requirement in facing the demands of the modern corporation. The second major finding was the central role organizational climate played, both as an intervening variable and as a predictor of perceived future success. Global climate is defined as the ambiance of an organization as reflected in its morale, conviviality, satisfaction and shared commitment. Nicholson and Brenner as well as Denison (1990) considered climate as essential for understanding organizational performance. A favourable attitudinal climate seems to be a precondition for the continued effectiveness of the high-performance, market-client driven organization. In the adapted model, the *long-term* criterion of survival was renamed survival and growth with emphasis on growth, and was complemented by market standing.

**Validity of non-financial performance measures**

In view of the potential of non-financial performance measures in creating value for the organization, the validity of these measures will be examined. Ittner and Larcker (1998a) reported on two research streams with regard to the association between non-financial performance measures and organizational performance. The first stream investigated claims that non-financial performance measures are “leading” indicators that provide information on future performance not reflected by traditional performance measures. However, research on the link between non-financial performance measures and financial performance produced mixed results. Three studies reported positive associations between customer satisfaction and financial performance (Anderson, Fornell & Lehmann, 1994; Banker, Potter & Srinivasan, 2000; Ittner & Larcker, 1996). Further research supported claims that customer satisfaction is a leading indicator of future consumer purchase behaviour, and growth in customers and current market values. However, organizational-level results varied by industry, with positive results in some industries and negative or insignificant results in others (Ittner & Larcker, 1998a, 1998b). The second stream of research focused on TQM and other advanced manufacturing systems. In nearly all studies positive correlations were found between the emphasis placed on interventions such as TQM and just-in-time (JIT) production practices or manufacturing flexibility, and non-financial measures such as defect rates, on-time delivery and machine utilization (Abernethy & Lillis, 1995; Banker, Potter & Schroeder, 1993; Daniel & Reitsperger, 1991a, 1991b; Perera, Harrison & Poole, 1997). Empirical support for the hypothesized performance benefits from these indicators is, however, marginal (Ittner & Larcker, 1998a).

The question arises why such a weak association exists between individual non-financial performance measures and the financial performance of organizations. Kaplan and Norton (2001) suggested two possible answers. Firstly, intangible assets provide value through organizational processes that convert intangible assets into customer and financial outcomes. It entails a non-linear process of value creation. The balance sheet, however, is a linear, additive document that records each class of asset separately and calculates the total by adding up each asset’s recorded value. Secondly, the value from intangible assets is indirect and, therefore, assets such as knowledge and technology seldom have a direct impact on revenue and profit. In fact, improvements in intangible assets affect financial performance through a series of cause-effect relationships that may comprise a number of stages (Becker & Huselid, 1998; Huselid, 1995).

## METHOD

### Sample

A non-probability sample of unit managers (units) was selected. The sample was selected by identifying part-time MBA students of the Graduate School of Business at the University of Stellenbosch who manage work units that meet the requirements of a work unit defined in the introduction to the paper. These students occupied full-time positions in middle management, senior management, and the professions. The unit sample comprises Modular English students of the 1998 intake, and all the Modular and Part Time students (English and Afrikaans groups) of the 1999 and 2000 intakes. Out of a possible number of 115 eligible work unit managers, 60 participated in the study. This figure represents a 52% participation that can be considered satisfactory.

The unit managers served as focal points for the identification of raters to make the 360° assessments of these managers' units since it is impossible to do a 360° rating on a unit without an appropriate focal point. Unit performance was rated by superiors, peers and subordinates of these unit managers, and a total of 257 completed Performance Index questionnaires were received at the time of the analysis. Although the objective initially was to obtain ratings from two subordinates, two peers and a single superior, the need for a large as possible sample size, in conjunction with the difficulties encountered when trying to apply a questionnaire of this length to respondents on this high a job level, necessitated a deviation from this ideal in a number of cases. The sample is presented in Table 1.

**TABLE 1**  
**SAMPLE OF WORK UNIT MANAGERS**

MBA Group	1998	1999	2000	Total
Modular English	6	7	15	28
Modular Afrikaans	-	12	11	23
Part Time	-	6	3	9
Total	6	25	29	60

### Developing the measurement instrument

The process for developing the Performance Index comprised three steps.

Step 1. Decide on a base-line structure for a model of work unit performance. Following a literature search covering organisational effectiveness, including performance measurement, a model of work unit performance was compiled. This entailed a synthesis of a systems model of organizational performance (Nicholson & Brenner, 1994), an adapted time-dimension model of organizational effectiveness (Gibson et al., 1991), and the outcomes of leadership (Conger & Kanungo, 1998). Since the underlying philosophy of this study is that work unit performance is the foremost criterion of leadership effectiveness, the dimensions of the performance construct conceptualized and operationalised here are intended to measure the outcomes of effective leader behaviour.

Step 2. Develop and verify appropriateness of the performance model and expand the dimensions of the model. The dimensions of Nicholson and Brenner's (1994) systems model were retained and expanded or adapted as follows:

1. Wealth, adaptability and climate, as well as future growth were retained.
2. Market share was expanded to address the needs of not-for-profit organizations and its name was changed to market standing.
3. Climate was split into work unit climate and individual climate (satisfaction) because a significant number of items

- reflected individual employee sentiments, including outcomes of leader effectiveness (Conger & Kanungo, 1998).
4. A short-term dimension termed production-efficiency (outputs) was added (Cockerill et al., 1993; Gibson et al., 1991).

In view of the main aim of step 2, a document that contained seven suggested dimensions of work unit effectiveness, with a summary description of items for each dimension, was sent to leadership associates of the Center for Leadership Studies for review and suggestions. The dimensions were production and efficiency, work unit climate, individual climate, capacity (wealth), market standing, adaptability, and future growth. Associates were specifically asked to obtain the co-operation of management peers in order to assist in the evaluation of the draft model. Two aspects were evaluated. The first aspect considered relevance of the dimensions to the conceptualization of the unit performance construct. A unit performance dimension was regarded as appropriate to the extent to which: The unit leader could be held accountable for the specific facets of unit performance, the unit leader directly affects the particular facet of unit performance, and top management considers the performance dimension as relevant. The second aspect that the panel evaluated was the comprehensiveness of the conceptualization of the model. Two questions were put to the panel: Should the model be expanded by adding additional effectiveness dimensions; and are there other issues, apart from the seven performance dimensions already included in the model, which top management might consider relevant when looking at unit performance?

Step 3. Consolidate the model and refine the questionnaire. The model was consolidated following a review of information from the field survey, and testing it against the literature. The only major adaptation made to the model was the inclusion of the dimension of core people processes. This dimension is based on Beckhard's (1969) and Beckhard and Harris's (1987) criteria of organizational health and effectiveness. Inclusion of this dimension was prompted partly by the positive field response to the importance of people-related dimensions such as adaptability and climate. Furthermore, important unit performance dimensions such as adaptability and climate do not materialize automatically, but require the effective functioning of core people processes. Conger and Kanungo (1989) recommended that, instead of traditional objective performance measures, leader effectiveness outcomes should rather be used, for example, "instrumental attitudes and behaviors that encourage the achievement of group objectives" (p. 39). It was decided not to include the medium term dimension of development as a separate dimension in the questionnaire but rather to subsume the developmental elements under the dimensions of capacity and core people processes.

An experiment in the area of judgement and decision-making research is relevant to scale construction of the PI. It was found that, when decision-makers are faced with both common and unique measures, they might place more weight on common measures than on unique measures (Slovic & MacPhillamy, 1974). This finding was repeated in a recent controlled study involving MBA students who evaluated two divisions of a clothing firm by means of the Balanced Scorecard. Experimental participants evaluated performance solely on the common measures (Lipe & Salterio, 2000). In view of this finding, emphasis on evaluation of both common and unique behavioural variables was stimulated by increasing understanding of non-traditional measures. This was achieved by including behavioral descriptions on all performance assessment scales. This would increase understanding of both the meaning of items as well as the relative strength of different points on each scale. Consequently, rating categories 5, 3, and 1 of the rating scale were anchored with behaviour descriptions. Caution with questionnaire construction and rater instruction was

considered to be essential for rating a diverse set of performance variables, particularly lesser known leading variables. The research copy of the Performance Index consisted of 56 questions covering eight dimensions. The dimensions, with a brief description of each dimension, are presented in Table 2. A five-point rating scale was used.

**TABLE 2**  
**BRIEF SUMMARIES OF PERFORMANCE DIMENSIONS**

1. Production and efficiency include quantitative outputs such as meeting goals, quantity, quality, cost-effectiveness, and task performance.
2. Core people processes reflect organizational effectiveness criteria such as goals and work plans, communication, organizational interaction, conflict management, productive clashing of ideas, integrity and uniqueness of the individual or group, learning through feedback, and rewarding performance.
3. Work unit climate is a global perception of the psychological environment of the unit, and gives an overall assessment of the integration, commitment and cohesion of the unit. It includes working atmosphere, teamwork, work group cohesion, agreement on core values and consensus regarding the vision, achievement-related attitudes and behaviours, and commitment to the unit.
4. Employee satisfaction centres around satisfaction with the task and work context, empowerment and career progress, as well as with outcomes of leadership, e.g. trust in and respect for the leader, and acceptance of the leader's influence.
5. Adaptability reflects the flexibility of the unit's management and administrative systems, core processes and structures, capability to develop new products/services, and versatility of staff and technology. Overall, it reflects the capacity of the unit to react appropriately and expeditiously to change.
6. Capacity (wealth of resources) reflects the internal strength of the unit, including financial resources, profits and investment; physical assets and materials supply; and quality and diversity of staff.
7. Market share/scope/standing includes market share (if applicable), competitiveness and market-directed diversity of products or services, customer satisfaction, and reputation for adding value to the organization.
8. Future growth serves as an overall index of projected future performance and includes profits and market share (if applicable), capital investment, staff levels and expansion of the unit.

## STATISTICAL ANALYSIS AND RESULTS

### Missing Values

Missing values presented a problem that had to be addressed before analysis could proceed. The classical treatment of the missing value problem through list-wise deletion of cases would, due to the extent of the problem and the length of the questionnaire, have dramatically reduced the sample size to 74, which would have made any meaningful statistical analysis impossible. Replacing the missing values with item means would also not have been advisable since it would effectively wash out most of the structure that exists in the data. Pair-wise deleting of cases also presented itself as a less than satisfactory alternative in that it results in a correlation matrix with extreme variation in N-values (a maximum of 256 and a minimum of 119 in this particular case). According to Jöreskog and Sörbom (1996a), experience indicates that such correlation matrices sometimes fail to be positive-definite. The possibility of using imputation to solve the missing value problem was subsequently explored. Imputation refers to a process of substituting of real values for missing values. The substitute values replaced for a case are derived from one or more other cases that have a similar response pattern over a set of matching variables (Jöreskog & Sörbom, 1996a). The ideal is to use matching variables that will not be utilized in the confirmatory factor analysis. This was, however, not possible in this case. The items least plagued by missing values were firstly identified. A set of eleven variables with three or less missing values per variable was subsequently defined to serve as matching variables. The PRELIS program (Jöreskog & Sörbom, 1996b) was used to

impute missing values. The subsequent PRELIS run on the reduced item set proved to be gratifyingly effective in countering the missing value problem. By default, cases with missing values after imputation are eliminated. After imputation, 241 cases with observations on all 56 items remained in the validation sample.

### Item Analysis

Each of the 8 PI sub-scales were subsequently item analyzed individually through the SPSS Reliability Procedure (SPSS, 1990) to identify and eliminate items not contributing to an internally consistent description of the unit performance facet in question. No items needed to be deleted based on the results of the item analyses. The results of the item analysis are shown in Table 3. Given the intended use of the PI as a comprehensive criterion measure against which to validate leadership and other competency assessments, the relatively high item homogeneity found for each sub-scale, as indicated by the Cronbach alpha values in Table 3, are extremely gratifying.

**TABLE 3**  
**RELIABILITY OF PI SUB-SCALE MEASURES**

Scale	Number of items	Alpha	Mean	Variance
Production & Efficiency	5	0,8310	18,6018	11,1563
Core People Processes	9	0,8843	30,8028	44,0485
Work Unit Climate	7	0,9036	24,8264	29,2312
Employee Satisfaction	9	0,9070	30,7043	45,3533
Adaptability	7	0,8449	24,0636	23,7768
Capacity	7	0,8510	22,2039	25,5409
Market Share	7	0,8856	24,2042	28,4474
Future Growth	5	0,8701	16,8019	16,5413

### Dimensionality Analysis

During the development of the PI, the objective was to construct essentially uni-dimensional sets of items to reflect variance in each of the eight latent variables collectively comprising the unit performance domain. Unrestricted principal component analyses with Varimax rotations were therefore performed on each on the eight PI sub-scales, each representing a facet of the multi-dimensional unit performance construct. The objective of these analyses was to confirm the uni-dimensionality of each sub-scale and to remove items with inadequate factor loadings and/or split heterogeneous sub-scales into two or more homogenous subsets of items if necessary (and make concomitant adjustments to the underlying unit performance model). The eigenvalue greater than unity rule of thumb was used to determine the number of factors to extract. SPSS (1990) was used for these analyses. A series of confirmatory factor analyses utilizing LISREL probably would have provided more stringent tests of the dimensionality of each sub-scale. Seven of the eight sub-scales passed the uni-dimensionality test. In the case of the employee satisfaction sub-scale two items [D3 & D4] had to be culled. The inclusion of these items in the specific sub-scale invariably resulted in the emergence of a second factor on which only the two items in question had a significant loading. No meaningful theme unique to the two items could be identified that distinguishes the second factor from the first factor. Culling these items from the respective sub-scales had the desired effect of reducing the dimensionality of the scale in which all remaining items had satisfactory ( $\lambda > 0, 64$ ) factor

loadings. The internal consistency of the scale remained essentially unaffected ( $\alpha = 0.9000$ ), although the mean and variance of the sub-scale distribution was necessarily lowered (25.3138 & 27.7204).

**Structural Equation Modelling**

Structural equation modelling (SEM) was used to perform a confirmatory factor analysis on the reduced data set obtained after imputation of missing values. Structural equation modelling on the PI in which each individual item serves as a manifest or indicator variable of the various latent leadership facets would, however, have resulted in a somewhat cumbersome and extensive exercise simply due to the number of items involved. Consequently two indicator variables were created from each sub-scale by calculating the unweighted average of the odd numbered items and the even numbered items of each scale. This should have the added advantage of creating more reliable indicator variables. The Cronbach alpha values reported in Table 3 should, however, not be simply extrapolated to the newly created composite indicator variables.

Each of the items comprising the PI was written so that a respondent's reaction to them would serve as a comprehensive and uncontaminated behavioural indicator of one of the facets of the multi-dimensional leadership domain. The conceptualization of the unit performance construct, in conjunction with the architecture of the PI, thus implies a specific factor structure or measurement model. The measurement model underlying the PI is shown in matrix format as equation 1.

$$X = \Lambda_X \xi + \delta \text{-----}1$$

Where X is a 16x1 column vector of observable indicator scores,  $\Lambda_X$  is a 16x8 matrix of factor loadings,  $\xi$  is a 28x1 column vector of latent leadership facets, and  $\delta$  is a 16x1 column vector of unique/measurement error components comprising the combined effect on X of systematic non-relevant influences and random measurement error (Jöreskog, 1993).

The measurement model implies two additional matrices. A symmetric 16x16 covariance/correlation matrix  $\Phi$  contains the correlations between the latent unit performance dimensions. A diagonal 16x16 matrix  $\theta\delta$  depicts the variance in the error terms associated with the indicator variables. The diagonal nature of the  $\theta\delta$  matrix implies that the error terms  $\delta_i$  are assumed to be uncorrelated across the indicator variables. If the measurement model would make provision for correlated error terms by freeing the off-diagonal elements of  $\theta\delta$ , it would imply the existence of additional common factors, not reflected in the model, but which also underly the responses to the indicator variables.

**Information on the parameters for the structural equation modelling analysis.**

LISREL 8.30 (Jöreskog, Sörbom, Du Toit & du Toit, 2000) was used to perform a confirmatory factor analysis on the PI to determine the fit of the model. The imputed data was first read into PRELIS (Jöreskog & Sörbom, 1996b) to compute a covariance matrix to serve as input for the LISREL analysis. Maximum likelihood estimation was used to estimate the parameters set free in the model. The latent variables contained in the model as such have no inherent scale, and neither are the values expressed in a meaningful unit of measurement. In specifying the model the scales of measurement of the latent variables were not specified by setting the factor loadings on the first observed variable to unity. Instead of defining the origin and unit of the latent variable scales in terms of observable reference variables, the latent variables were rather standardized (Jöreskog & Sörbom, 1993). The unit of measurement

thus becomes the standard deviation  $\sigma_i[\xi]$ . All factor loadings of each latent unit performance variable were set free to be estimated, but only with regards to its designated observed variables. All remaining elements of  $\Lambda_X$  were fixed at zero loadings to reflect the assumed factorial simplicity of the PI items (Tabachnick & Fidell, 1989). The elements of  $\Phi$  and the diagonal elements of  $\theta\delta$  were treated by default as free.

**An assessment of model fit through multiple fit indices**

An admissible final solution of parameter estimates was obtained after 9 iterations. The full spectrum of indices provided by LISREL to assess the absolute and comparative fit of the model is shown in Table 4.

**TABLE 4  
GOODNESS OF FIT STATISTICS**

---

Degrees of Freedom = 76
Minimum Fit Function Chi-Square = 131,30 (P = 0,00)
Normal Theory Weighted Least Squares Chi-Square = 129,67 (P = 0,00012)
Estimated Non-centrality Parameter (NCP) = 53,67
90 Percent Confidence Interval for NCP = (26,05 ; 89,17)
Minimum Fit Function Value = 0,55
Population Discrepancy Function Value (F0) = 0,22
90 Percent Confidence Interval for F0 = (0,11 ; 0,37)
Root Mean Square Error of Approximation (RMSEA) = 0,054
90 Percent Confidence Interval for RMSEA = (0,038 ; 0,070)
P-Value for Test of Close Fit (RMSEA < 0,05) = 0,32
Expected Cross-Validation Index (ECVI) = 1,04
90 Percent Confidence Interval for ECVI = (0,93 ; 1,19)
ECVI for Saturated Model = 1,13
ECVI for Independence Model = 13,60
Chi-Square for Independence Model with 120 Degrees of Freedom = 3231,43
Independence AIC = 3263,43
Model AIC = 249,67
Saturated AIC = 272,00
Independence CAIC = 3335,18
Model CAIC = 518,76
Saturated CAIC = 881,93
Normed Fit Index (NFI) = 0,96
Non-Normed Fit Index (NNFI) = 0,97
Parsimony Normed Fit Index (PNFI) = 0,61
Comparative Fit Index (CFI) = 0,98
Incremental Fit Index (IFI) = 0,98
Relative Fit Index (RFI) = 0,94
Critical N (CN) = 197,65
Root Mean Square Residual (RMR) = 0,013
Standardized RMR = 0,025
Goodness of Fit Index (GFI) = 0,94
Adjusted Goodness of Fit Index (AGFI) = 0,89
Parsimony Goodness of Fit Index (PGFI) = 0,52

---

The  $\chi^2$  test statistic tests the null hypothesis, shown as equation 2, that the population covariance matrix is equal to the population covariance matrix implied by the model.

$$H_0: \Sigma = \Sigma(\theta) \text{-----}2$$

The p-value (0.00012) associated with the  $\chi^2$  value in Table 4 clearly indicates a highly significant test statistic. However the  $\chi^2$  measure is only distributed asymptotically as a  $\chi^2$  distribution. This causes the frustrating dilemma that just at the point where the distributional assumption of the test statistic becomes tenable the statistical power of the

test also becomes extremely high. Given the sample size involved in this study it therefore seems somewhat premature to conclude poor model fit based on the large and significant  $\chi^2$  alone.

Expressing the  $\chi^2$  value in terms of its degrees of freedom has been suggested as a way of getting round the aforementioned problems associated with this measure. This is not routinely provided by LISREL as part of its repertoire of fit measures and thus not shown in Table 4. A value of 1.71 results in this case. The interpretation of the ratio  $\chi^2/df$ , however, seems somewhat problematic in that no clear, generally agreed upon guidelines seem to exist (Kelloway, 1998; Medsker, Williams & Holahan, 1994). Ratios between 2 and 5 seem to be regarded as indicative of good fit. Ratios less than 2 have, however, been interpreted as indicating over-fitting. Judged by these standards the model could, when viewed optimistically, be seen to fit the data well, or, when viewed somewhat more pessimistically, be seen to have been over-fitted. Kelloway (1998) comments that the aforementioned interpretative standards have very little justification other than researcher's personal modelling experience and advises against a strong reliance on its use.

The root mean squared residual (RMR) and standardized RMR reflect the square root of the mean squared difference between the observed and estimated covariance matrices. Table 4 reports values of 0,013 and 0,025 for these two measures of fit. Values of less than 0,05 on the latter index are regarded as indicative of a model that fits the data well (Kelloway, 1998). The root mean squared error of approximation (RMSEA) expresses the difference between the observed and estimated covariance matrices in terms of the degrees of freedom of the model (Steiger, 1990). Though rarely encountered, RMSEA values below 0,01 would indicate a model that fits the data exceptionally well. RMSEA values below 0, 05 would indicate a very good fit and values below 0,1 a good fit (Kelloway, 1998; Steiger, 1990). Hair, Anderson, Tatham and Black (1995) consider RMSEA values between 0,05 and 0,08 indicative of acceptable fit. Brown and Cudeck (1993) regard a RMSEA value of 0,05 indicative of a close fit and RMSEA values up to 0,08 indicative of reasonable errors of approximation. A RMSEA value of 0,054 results in Table 4. The 90% confidence interval for RMSEA shown in Table 3 (0,038 - 0,070) indicates that the fit of the PI measurement model could be regarded as good to very good. A test of the significance of the obtained value is performed by LISREL by testing  $H_0: RMSEA \leq 0,05$  against  $H_a: RMSEA > 0,05$ . Table 4 indicates that the obtained RMSEA value of 0,054 is not significantly smaller than the target value of 0, 05 (i.e.  $H_0$  is not rejected;  $p > 0,05$ ) and since the confidence interval does include the target value of 0,05, a good or acceptable fit seems to have been achieved. This conclusion is supported by the aforementioned Standardized RMR value of 0,025.

The goodness-of-fit index (GFI) measures are "based on a ratio of the sum of the squared discrepancies to the observed variances (for generalized least squares, the maximum likelihood version is somewhat more complicated)" (Kelloway, 1998, p. 27). The adjusted GFI (AGFI) adjusts the GFI for degrees of freedom in the model (Kelloway, 1998). Both these two measures should be between zero (poor fit) and unity (perfect fit) (Jöreskog & Sörbom, 1993) with values exceeding 0,9 indicating good fit to the data (Kelloway, 1998). Evaluating the fit of the model in terms of these two indices (0,94 & 0, 89) a relatively favourable conclusion on model fit emerges. Kelloway (1998), however, warns that these guidelines for the interpretation of GFI and AGFI are grounded in experience, are somewhat arbitrary and should therefore be used with some circumspection.

**An examination of the obtained solution**

All estimated factor loadings  $\lambda_{ij}$  in  $\Lambda_X$  shown in Table 5 differ significantly ( $p < 0, 05$ ) from zero (standard errors and t-values are not shown). The fit of the model should therefore deteriorate significantly if any of the existing paths in the measurement model would be pruned away by fixing the corresponding parameters in  $\Lambda_X$  to zero and thus effectively eliminating the item parcel in question from their current sub-scales.

**TABLE 5**  
**COMPLETELY STANDARDIZED LAMBDA-X, FACTOR-LOADING MATRIX**

	PROD	CORE	CLIM	SATIS	ADAP	CAPA	MARK	GROW
PROD_1	0,91							
PROD_2	0,73							
CORE_1		0,91						
CORE_2		0,89						
CLIM_1			0,95					
CLIM_2			0,92					
SATIS_1				0,94				
SATIS_2				0,89				
ADAP_1					0,89			
ADAP_2					0,87			
CAPA_1						0,78		
CAPA_2						0,87		
MARK_1							0,88	
MARK_2							0,86	
GROW_1								0,80
GROW_2								0,80

Two questions, however, remain. Although all item parcels appear to significantly reflect the unit performance dimension it was designed to denote, the question firstly arises how well each item parcel measures its designated unit performance dimension. The proportion of item parcel variance that is explained by the model is shown in Table 6 for each of the PI indicator variables. The values shown in Table 6 should simultaneously be interpreted as lower bound estimates of the item parcel reliabilities  $\rho_{ii}$ .  $\theta\delta_i$  represent the error variance elements of the completely standardized diagonal matrix  $\theta\delta$  (shown in Table 7. Since the error term  $\delta_i$  comprise not only a true random measurement error component but also a systematic error component unique to  $X_i$ ,  $\rho_{ii}$  can also be interpreted as an indicator variable validity coefficient expressing the success with which the latent unit performance dimension  $\xi_j$  manifests itself in the indicator variable  $X_{ij}$ . Inspection of Tables 5, 6 and 7 indicates that the majority of item parcels do provide relatively uncontaminated and comprehensive reflections of their designated latent dimensions. The second item parcel of the Production and Efficiency sub-scale (Prod\_2) is the only indicator that appears to have somewhat questionable relevance for the unit performance dimension to which it is currently linked with. Only approximately 53% of the variance in Prod\_2 can be explained in terms of the first latent variable while the remaining approximately 47% of the variance in this item parcel should be attributed to measurement error.

**TABLE 6**  
**SQUARED MULTIPLE CORRELATIONS FOR ITEM PARCELS**

PROD_1	PROD_2	CORE_1	CORE_2	CLIM_1	CLIM_2	SATIS_1	SATIS_2
0,83	0,53	0,83	0,78	0,91	0,84	0,88	0,79
ADAP_1	ADAP_2	CAPA_1	CAPA_2	MARK_1	MARK_2	GROE_1	GROW_2
0,79	0,76	0,61	0,76	0,77	0,73	0,64	0,65

**TABLE 7**  
**COMPLETELY STANDARDIZED THETA-DELTA MATRIX**

PROD_1	PROD_2	CORE_1	CORE_2	CLIM_1	CLIM_2	SATIS_1	SATIS_2
0,17	0,47	0,17	0,22	0,09	0,16	0,12	0,21
ADAP_1	ADAP_2	CAPA_1	CAPA_2	MARK_1	MARK_2	GROE_1	GROW_2
0,21	0,24	0,39	0,24	0,23	0,27	0,36	0,35

The phi-matrix of correlations between the 8 latent unit performance dimensions is shown in Table 8. The off-diagonal elements of the  $\Phi$ -matrix are the inter-unit performance dimension correlations disattenuated for measurement error. The correlations are all moderate to high thus suggesting the need to expand the model through the addition of a limited set of two, or possibly three, second-order factors. This is to a certain extent to be expected given firstly the inter-dependent nature of the dimensions in the model and secondly, the time dimension underlying the model. The  $\Phi$ -matrix is, however, fortunately still positive definite with no off-diagonal entries exceeding unity. Experience seems to suggest that for oblique measurement models with a reasonable number of latent variables, such as the PI model, there always exists the real danger that the resultant phi-matrix could not be positive definite. This would have severely complicated further investigations of the higher-order structure of the instrument.

**TABLE 8**  
**COMPLETELY STANDARDIZED PHI MATRIX OF DISATTENUATED OF UNIT PERFORMANCE DIMENSION CORRELATIONS**

	PROD	CORE	CLIM	SATIS	ADAP	CAPA	MARK	GROW
PROD	1,00							
CORE	0,66	1,00						
CLIM	0,66	0,83	1,00					
SATIS	0,68	0,81	0,83	1,00				
ADAP	0,52	0,74	0,68	0,72	1,00			
CAPA	0,53	0,72	0,66	0,66	0,86	1,00		
MARK	0,61	0,71	0,70	0,66	0,79	0,88	1,00	
GROW	0,54	0,66	0,60	0,57	0,76	0,84	0,83	1,00

Table 8 seems to suggest possibly two second-order factors with Employee Satisfaction, Climate and Core People Processes loading on the first higher-order factor and Future Growth, Market Standing, Capacity, and Adaptability loading on the second higher-order factor. The first second-order factor could possibly be described as a "softer" people performance factor, while the second in contrast could possibly be viewed as a "harder" systems performance factor. Production and Efficiency, the classical measure of unit performance, though definitely not unrelated to the previous two clusters,

nonetheless seems to resist fusion into a more general latent performance dimension. This hypothetical higher-order factor structure, or some variant of it, will subsequently be confronted with the same data set via second-order confirmatory factor analysis, utilizing LISREL.

#### Examination of residuals

Residuals refer to the differences between corresponding cells in the observed and fitted covariance/correlation matrices (Jöreskog & Sörbom, 1993). Residuals, and especially standardized residuals, thereby provide diagnostic information on sources of lack of fit in models (Jöreskog & Sörbom, 1993; Kelloway, 1998). The stem-and-leaf plot of standardized residuals (not shown) indicates that the standardised residuals are distributed approximately normal. The leptocurtic nature of the distribution would suggest that relatively few covariance terms in the observed covariance matrix were inadequately accounted for by the fitted model. In addition, the largest and smallest standardized residuals seem to be of only modest magnitude ( $\pm |2,8|$ ). Good model fit is further indicated by the fact that the standardized residuals for all pairs of observed variables approximately fall on a 45° straight line in the Q-plot (not shown).

#### Model modification indices

Although all item parcels appear to significantly reflect the unit performance dimension it was designed to denote, the question furthermore arises whether the addition of more paths to the model would significantly improve the parsimonious fit of the model. Examination of the modification indices and the completely standardized expected parameter change associated with the fixed parameters in  $\Delta\chi^2$ , indicates that no paths, if added to the model, should result in notable decreases in the  $\chi^2$  measure. Examination of the modification indices and the completely standardized expected parameter change associated with the fixed parameters in the  $\Theta\delta$  matrix reveal no covariance terms that, if set free, would result in major decreases in the  $\chi^2$  measure. Only three covariance terms are isolated in the interactive path diagram as possible  $\Theta\delta$  parameters to be freed. The expected magnitude of the covariate estimates (less than 0,05), however, hardly warrants seriously considering setting these three parameters free. This in turn would suggest that the assumption of uncorrelated error terms remain largely tenable.

## DISCUSSION

Organizational units exist with the explicit purpose to produce a specific product or service. Traditionally organisational units were therefore evaluated in terms of the efficiency with which they fulfil the objective for which they exist, and in terms of the extent to which they satisfy their client's quality, quantity and distributional expectations. If an organizational unit consistently succeeds in delivering a superior output to its clients over an extended period of time, it thereby develops an elevated market standing. It develops a known high reputation and tends to become synonymous with the type of product/service in question. It simultaneously expands its market share. If an organizational unit currently has a high market standing due to its consistently efficient delivery of a superior product or service and the organizational unit has the ability to adapt to internal and/or external environmental changes, the unit will currently be characterized by high future growth prospects.

Current high market standing due to consistently efficient delivery of a superior product/service however, cannot be achieved without at least three broad prerequisites being met. Efficient core people processes and structures represent a first, indispensable requirement for high unit production efficiency. A smooth-running, quick-response, low-friction,



high-energy human system must be available to pursue the production objectives. Continuous and sufficient access to superior quality physical, financial, natural and human resources represents a second, equally indispensable requirement to achieve high production efficiency. An organizational unit cannot by itself successfully activate its human system in pursuit of its production objectives. A work unit climate that constitutes an expression of a set of shared core values and a commitment to a shared unit vision and mission represents a third essential requirement to achieve high production efficiency.

Employee satisfaction represents a barometer of the extent to which an organizational unit succeeds in establishing the prerequisites for efficient and superior production. Employee satisfaction, in metaphorical sense, therefore, represents the human resource equivalent of the caged mining canary in days gone by.

Assessments of organizational unit performance should undoubtedly include a measure of production efficiency. However, to restrict the assessment of unit performance to this dimension only, would, in terms of the foregoing argument, result in a deficient criterion measure. Efficient/successful units, over time, develop a high market standing and continuously impress investors with their future growth prospects. To achieve this, however, requires efficient core people processes, capacity and a supportive climate. All eight aspects of the PI thus need to form part of the spectrum of unit performance dimensions that should be assessed when evaluating the success of an organizational unit. Effective functioning of units requires regular appraisal of the various dimensions that constitute unit performance. The Performance Index was developed to identify areas where remedial interventions are required. The actual use of the Performance Index for this purpose, however, requires evidence that the Performance Index does in fact provide a valid and reliable measure of the unit performance construct.

To be able to conduct meaningful research in critical areas such as, for example, leadership and Performance Management – which are the two major research areas of the Centre for Leadership Studies – the availability of a credible criterion measure is essential. In the absence of an appropriate unit performance measure, no claim can be made that the models underlying the instruments developed at the Centre, namely the Leadership Behaviour Inventory (LBI) and the Performance Management Audit Questionnaire constitute prerequisites for ultimate effective unit performance. The existing work unit performance measures do not seem to encompass all the performance measures for which the unit leader could be held accountable, or on which Performance Management interventions should impact (Cockerill et al., 1993; Nicholson & Brenner, 1994). In addition, a measure is required that would be applicable across various units within a single organisation, and across different organisations and industries. The Performance Index was developed for exactly this purpose. The Performance Index provides a more comprehensive measurement of leading and lagging performance measures than the comparable Unit Performance Questionnaire (UPQ) (Cockerill et al., 1993) discussed earlier. Specifically, the Performance Index expands on the UPQ by the inclusion of dimensions such as core people processes, market share/scope/standing and future growth, and by covering existing dimensions more comprehensively.

The Performance Index was developed to fulfil the aforementioned need for a diagnostic and applied research instrument. The results reported here on the item analyses and the confirmatory factor analysis provides reasonable psychometric support for using the PI for this purpose. The possible existence of a higher-order factor structure should,

however, be explored by defining a set of second-order factors and confronting the resultant model in a further confirmatory analysis with the present data set. The integrity of the phi matrix fortunately makes this a feasible analysis. Should the analysis on the present data set support the postulated higher-order factor structure, both the primary and the extended measurement models should be confronted with a fresh data set. The unit performance model depicted as an exogenous measurement model should in addition, given the foregoing argument, be expanded into a fully-fledged structural model that captures the inter-relationships between the eight unit performance latent variables. This expanded, more complex model should also be confronted with the current data set and if corroborated, also with a fresh data set.

The responsibility for the performance of any organizational unit on these eight performance dimensions ultimately lies with the leadership of the unit. Leadership in this sense constitutes a complex process expressing itself in an array of inter-dependent behavioural actions and driven by an intricate nomological network of situational and person-centred latent variables. The process essentially entails (a) the assessment of the internal and external environment of the unit, (b) the development and selling of an environmentally appropriate yet challenging vision for the unit, (c) the preparation of the unit for the implementation of the vision, and finally (d) the bold yet honest implementation of the vision by continually monitoring, revitalizing, fine-tuning and orchestrating a multitude of prerequisites for unit success in terms of the vision. Given the perceived pivotal role of leadership in organizational unit performance, the nature of the presumed relationship should be captured in a comprehensive structural model that would explain the manner in which the various latent leadership dimensions affect the endogenous unit performance latent variables. The evidence on the validity of the measurement model underlying the PI reported here, in conjunction with the results on the LBI reported in Spangenberg and Theron (2002), now paves the way for proceeding with the rather daunting task of explicating and evaluating such a comprehensive leadership-unit performance structural model.

## REFERENCES

- Abernethy, M.A. & Lillis, A.M. (1995). The impact of manufacturing flexibility on management control system design. *Accounting, Organizations and Society*, 20 (4), 241–258.
- Anderson, E.W., Fornell, C. & Lehmann, L.R. (1994). Customer satisfaction, market share, and profitability: Findings from Sweden. *Journal of Marketing Research*, 58 (July), 53–66.
- Balkcom, J.E., Ittner, C.D. & Larcker, D.F. (1997). Strategic performance measurement: Lessons learned and future directions. *Journal of Strategic Performance Measurement*, 1 (2), 22–32.
- Banker, R.D., Field, J.M., Schroeder, R.G. & Sinha, K.K. (1996). Impact of work teams on manufacturing performance: A longitudinal field study. *Academy of Management Journal*, 39 (4), 867–890.
- Banker, R., Potter, G. & Schroeder, R.G. (1993). Reporting manufacturing performance measures to workers: An empirical study. *Journal of Accounting Research*, 5 (Fall), 3–55.
- Banker, R., Potter, G. & Srinivasan, D. (2000). An empirical investigation of an incentive plan that includes non-financial performance measures. *The Accounting Review*, 75 (1), 65–92.
- Becker, B. & Huselid, M. (1998). *High performance work systems and firm performance: A synthesis of research and managerial implications*. In *Research in Personnel and Human Resources Management* (53–101). Greenwich, CT: JAI Press.
- Beckhard, R. (1969). *Organization development: Strategies and models*. Reading, Mass: Addison Wesley.

- Beckhard, R. & Harris, R.T. (1987). *Organizational transitions: Managing complex change*. Reading, Mass: Addison Wesley.
- Brancato, C.K. (1995). *New performance measures: A research report*. Report no. 1118 - 95 - RR. The Conference Board, New York.
- Brown, M.W. & Cudeck, R. (1993). Alternative ways of assessing model fit. In K.A. Bollen & J.S. Long (Eds), *Testing Structural Equation Models*. Newbury Park: Sage Publications.
- Cameron, K.S. (1986). Effectiveness as paradox: Consensus and conflicts in conceptions of organizational effectiveness. *Management Science*, 32, 539-553.
- Cameron K.S. & Whetton, D.A. (1983). Organizational effectiveness: One model or several? In K.S. Cameron & D.A. Whetton, (Eds.), *Organizational effectiveness: A comparison of multiple models* (pp.1-24). New York: Academic Press.
- Cockerill, A.P., Schroder, H.M. & Hunt, J.W. (1993). *Validation study into the High Performance Managerial Competencies*. Unpublished report, London Business School, London.
- Conger, J.A. & Kanungo, R.N. (1998). *Charismatic leadership in organizations*. London: Sage Publications.
- Daniel, S.J. & Reitsperger, W.D. (1991a). Linking quality strategy with management control systems: Empirical evidence from Japanese industry. *Accounting, Organizations and Society*, 16 (7), 601-618.
- Daniel, S.J. & Reitsperger, W.D. (1991b). Management control systems for J.I.T.: An empirical comparison of Japan and the U.S. *Journal of International Business Studies*, 22 (4), 603-607.
- Delaney, J.T. & Huselid, M.A. (1996). The impact of human resource management practices on perceptions of organizational performance. *Academy of Management Journal*, (39) 4, 949-969.
- Denison, D.R. (1990). *Corporate culture and organizational effectiveness*. New York: Wiley.
- Eccles, R.G. (1991). The performance measurement manifesto. *Harvard Business Review*, 69 (2), 131-137.
- Fisher. (1995). Use of non-financial performance measures. In S.M. Young (Ed.), *Readings in Management Accounting* (329-335). Englewood Cliffs, NJ: Prentice Hall.
- Gibson, J.L., Ivancevich, J.M. & Donnelly, J.H. (1991). *Organizations*. Irwin: Boston.
- Hair, J.F., Anderson, R.E., Tatham, R.L. & Black, W.C. (1995). *Multivariate data analysis with readings*. New Jersey: Prentice-Hall, Inc.
- Huang, T. (1997). The effect of participative management on organizational performance: The case of Taiwan. *The International Journal of Human Resource Management*. 8 (5), 677-689.
- Huselid, M.A. (1995). The impact of human resource management practices on turnover, productivity, and corporate financial performance. *Academy of Management Journal*, 38, 635-672.
- Ittner, C.D. & Larcker, D.F. (1996). Measuring the impact of quality initiatives on firm financial performance. *Advances in the Management of Organizational Quality*, 1, 1-37.
- Ittner, C.D. & Larcker, D.F. (1998a). Innovations in performance measurement: Trends and research implications. *Journal of Management Accounting Research*, 10, 205-239.
- Ittner, C.D. & Larcker, D.F. (1998b). Are non-financial measures leading indicators of financial performance? An analysis of customer satisfaction. *Journal of Accounting Research*, 36, 1-35.
- Jöreskog, K.G. (1993). Testing structural equation models. In Bollen, K.A. & Long, J.S. *Testing Structural Equation Models*. Newbury Park: Sage Publications.
- Jöreskog, K.G. & Sörbom, D. (1993). *LISREL 8: Structural Equation Modeling with SIMPLIS command language*. Chicago: Scientific Software International.
- Jöreskog, K.G. & Sörbom, D. (1996a). *LISREL 8: User's reference guide*. Chicago: Scientific Software International.
- Jöreskog, K.G. & Sörbom, D. (1996b). *PRELIS 2: User's reference guide*. Chicago: Scientific Software International.
- Jöreskog, K.G., Sörbom, D., Du Toit, S. & du Toit, M. (2000). *LISREL 8: New statistical features*. Chicago: Scientific Software International.
- Kaplan, R.S. & Norton, D.P. (2001). Transforming the Balanced Scorecard from performance measurement to strategic management: Part 1. *Accounting Horizons*, 15 (1), 87-102.
- Kelloway, E.K. (1998). *Using LISREL for structural equation modeling: a researcher's guide*. Thousand Oaks: Sage Publications, Inc.
- Lewin, A.Y. & Minton, J.W. (1986). Determining organizational effectiveness: Another look, and an agenda for research. *Management Science*, 32, 514-538.
- Lipe, M. & Salterio, S. (2000). The Balanced Scorecard: Judgmental effects of common and unique performance measures. *The Accounting Review*, 75 (3), 283-298.
- Medsker, G.J., Williams, L.J. & Holahan, P.J. (1994). A review of current practices for evaluating causal models in organizational behavior and human resources management research. *Journal of Management*, 20, 439-464.
- Miles, R.E. & Snow, C.C. (1978). *Organizational strategy, structure and processes*. New York: McGraw-Hill.
- Miles, R.H. (1980). *Macro-organizational behaviour*. Glenview, IL: Scott, Foresman.
- Mohr, L.B. (1983). The Implications of effectiveness theory for managerial practice in the public sector. In K.S. Cameron & D.A. Whetton (Eds.), *Organizational effectiveness: A comparison of multiple models* (225-239). New York: Academic Press.
- Nicholson, N. & Brenner, S.O. (1994). Dimensions of perceived organizational performance: Tests of a model. *Applied Psychology: an International Review*, 3 (1), 69-108.
- Nunnally, J.C. (1978). *Psychometric theory*. McGraw-Hill: New York.
- Perera, S., Harrison, G. & Poole, M. (1997). Customer-focused manufacturing strategy and the use of operations-based non-financial performance measures: A research note. *Organizations and Society*, 15 (3), 199-220.
- Quinn, R.E. & Rohrbaugh, J.A. (1983). A spatial model of effectiveness criteria: Towards a competing values approach to organizational analysis. *Management Science*, 29 (3), 363-377.
- Rucci, A.J., Kirn, S.P. & Quinn, R.T. (1998). The employee-customer profit chain at Sears. *Harvard Business Review*, 76 (1), 82-97.
- Schroder, H.M. & Cockerill, A.P. (1992). *Development of a multi-rater competence utilization questionnaire*. Unpublished report, University of South Florida, Tampa.
- Seashore, S.E. (1983). A framework for an integrated model of organizational effectiveness. In K.S. Cameron & D.A. Whetton (Eds.), *Organizational effectiveness: A comparison of multiple models*. San Diego: Academic Press.
- Slovic, P. & MacPhillamy, D. (1974). Dimensional commensurability and cue utilization in comparative judgment. *Organizational Behaviour and Human Performance*, 11, 172-194.
- Spangenberg, H.H., Schroder, H.M., Duvenage, A. & Theron, C.C. (1999). *Adapting a leadership competence utilization questionnaire for South African managers (Phases 1 and 2)*. Paper presented at the Sixth European Congress of Psychology, Rome.
- Spangenberg, H.H. & Theron, C.C. (2002). Development of a uniquely South African leadership development questionnaire. *South African Journal of Psychology*, 32 (2), 9-25.
- SPSS. (1990). *SPSS reference guide*. Chicago: SPSS International.
- Steiger J.H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*, 25 (2), 173-180.
- Tabachnick, B.G. & Fidell, L.S. (1989). *Using multivariate statistics* (second edition). New York: Harper Collins Publishers.
- Youndt, M.A., Snell, S.A., Dean, J.W. & Lepak, D.P. (1996). Human resource management, manufacturing strategy, and firm performance. *Academy of Management Journal*, 39 (4), 836-866.